

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 7, 2001, 10:41:29 ; Search time 1719.24 Seconds

(without alignments)  
8257.465 Million cell updates/sec

Title: US-09-376-633-1

Perfect score: 1029

Sequence: 1 ATGGATATATGCACTTCAACT.....TAACCTTACGAAAGAGTCA 1029

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 13168883 seqs, 6898233199 residues

Total number of hits satisfying chosen parameters: 26337766

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: Listing first 45 summaries

Pending Patents\_NA\_Main:\*

1: /cgn2\_6/ptodata/2/pna/US081\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/pna/US06\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/pna/US07\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/pna/US080\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/pna/US081\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/pna/US082\_COMB.seq.\*  
7: /cgn2\_6/ptodata/2/pna/US083\_COMB.seq.\*  
8: /cgn2\_6/ptodata/2/pna/US084\_COMB.seq.\*  
9: /cgn2\_6/ptodata/2/pna/US085\_COMB.seq.\*  
10: /cgn2\_6/ptodata/2/pna/US086\_COMB.seq.\*  
11: /cgn2\_6/ptodata/2/pna/US087\_COMB.seq.\*  
12: /cgn2\_6/ptodata/2/pna/US088\_COMB.seq.\*  
13: /cgn2\_6/ptodata/2/pna/US089\_COMB.seq.\*  
14: /cgn2\_6/ptodata/2/pna/US090\_COMB.seq.\*  
15: /cgn2\_6/ptodata/2/pna/US091\_COMB.seq.\*  
16: /cgn2\_6/ptodata/2/pna/US092\_COMB.seq.\*  
17: /cgn2\_6/ptodata/2/pna/US093\_COMB.seq.\*  
18: /cgn2\_6/ptodata/2/pna/US094\_COMB.seq.\*  
19: /cgn2\_6/ptodata/2/pna/US095A\_COMB.seq.\*  
20: /cgn2\_6/ptodata/2/pna/US095B\_COMB.seq.\*  
21: /cgn2\_6/ptodata/2/pna/US095C\_COMB.seq.\*  
22: /cgn2\_6/ptodata/2/pna/US095D\_COMB.seq.\*  
23: /cgn2\_6/ptodata/2/pna/US096A\_COMB.seq.\*  
24: /cgn2\_6/ptodata/2/pna/US096B\_COMB.seq.\*  
25: /cgn2\_6/ptodata/2/pna/US096C\_COMB.seq.\*  
26: /cgn2\_6/ptodata/2/pna/US096E\_COMB.seq.\*  
27: /cgn2\_6/ptodata/2/pna/US096F\_COMB.seq.\*  
28: /cgn2\_6/ptodata/2/pna/US097A\_COMB.seq.\*  
29: /cgn2\_6/ptodata/2/pna/US097B\_COMB.seq.\*  
30: /cgn2\_6/ptodata/2/pna/US097C\_COMB.seq.\*  
31: /cgn2\_6/ptodata/2/pna/US098\_COMB.seq.\*  
32: /cgn2\_6/ptodata/2/pna/US098A\_COMB.seq.\*  
33: /cgn2\_6/ptodata/2/pna/US098B\_COMB.seq.\*  
34: /cgn2\_6/ptodata/2/pna/US098C\_COMB.seq.\*  
35: /cgn2\_6/ptodata/2/pna/US098D\_COMB.seq.\*  
36: /cgn2\_6/ptodata/2/pna/US098E\_COMB.seq.\*  
37: /cgn2\_6/ptodata/2/pna/US098F\_COMB.seq.\*  
38: /cgn2\_6/ptodata/2/pna/US098G\_COMB.seq.\*  
39: /cgn2\_6/ptodata/2/pna/US098H\_COMB.seq.\*  
40: /cgn2\_6/ptodata/2/pna/US098I\_COMB.seq.\*  
41: /cgn2\_6/ptodata/2/pna/US098J\_COMB.seq.\*  
42: /cgn2\_6/ptodata/2/pna/US098K\_COMB.seq.\*  
43: /cgn2\_6/ptodata/2/pna/US098L\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1029	100.0	1029	1 PCT-US98-24852-1	Sequence 1, Appl 1
2	1029	100.0	1029	17 US-09-369-760-1	Sequence 1, Appl 1
3	1029	100.0	1029	17 US-09-376-633-1	Sequence 1, Appl 1
4	1029	100.0	3336	1 PCT-US98-24852-2	Sequence 7, Appl 1
5	1029	100.0	3336	1 PCT-US98-24852-3	Sequence 7, Appl 1
6	1029	100.0	3336	1 PCT-US98-24852-4	Sequence 7, Appl 1
7	1029	100.0	3336	1 PCT-US98-24852-5	Sequence 7, Appl 1
8	1029	100.0	3336	1 PCT-US98-24852-6	Sequence 7, Appl 1
9	1029	100.0	3336	13 US-08-977-553-5	Sequence 7, Appl 1
10	1029	100.0	3336	16 US-08-977-866-7	Sequence 7, Appl 1
11	1029	100.0	3336	16 US-09-227-806-7	Sequence 7, Appl 1
12	1029	100.0	3336	17 US-09-369-700-7	Sequence 7, Appl 1
13	1029	100.0	3336	17 US-09-376-633-7	Sequence 7, Appl 1
14	1029	100.0	3336	18 US-09-492-479-7	Sequence 7, Appl 1
15	1029	100.0	3336	18 US-08-827-356-1376	Sequence 1376, Ap
16	1029	100.0	3336	13 US-08-611-529-7446	Sequence 7446, Ap
17	1029	100.0	3336	23 US-08-831-156A-107	Sequence 107, App
18	1029	100.0	3336	12 US-08-781-986A-142	Sequence 142, App
19	1029	100.0	3336	13 US-08-956-171-142	Sequence 142, App
20	1029	100.0	3336	13 US-08-956-171B-142	Sequence 142, App
21	985.6	95.8	7003	35 US-60-038-697-616	Sequence 616, App
22	985.6	95.8	7003	36 US-60-046-714-618	Sequence 618, App
23	515.2	50.1	1053	18 US-09-750-969-2343	Sequence 2343, Ap
24	515.2	50.1	8374	38 US-60-068-228-315	Sequence 315, App
25	515.2	50.1	8413	35 US-60-038-081-306	Sequence 306, App
26	515.2	50.1	8414	36 US-60-046-933-313	Sequence 313, App
27	494.6	48.1	509	24 US-09-620-608-372	Sequence 372, App
28	494.6	48.1	509	46 US-60-144-883-372	Sequence 372, App
29	352.2	34.2	1269	1 PCT-US97-02318-946	Sequence 946, App
30	352.2	34.2	1269	1 PCT-US98-24857-3	Sequence 946, App
31	352.2	34.2	1269	13 US-08-903-470-946	Sequence 946, App
32	352.2	34.2	1269	17 US-09-369-700-3	Sequence 946, App
33	352.2	34.2	1269	17 US-09-376-633-3	Sequence 946, App
34	151.8	14.8	2507	26 US-09-663-779-1376	Sequence 1376, Ap
35	143	13.9	258	26 US-09-663-779-829	Sequence 829, App
36	120.4	11.7	5567	12 US-08-899-241-1	Sequence 2251, Ap
37	115.4	11.2	1101	17 US-09-328-352-1336	Sequence 1336, Ap
38	114.2	11.1	1101	22 US-09-583-110-2144	Sequence 2144, Ap
39	114.2	11.1	1206	15 US-09-107-433-2289	Sequence 2289, Ap
40	113.8	11.1	14273	13 US-08-961-527-40	Sequence 40, Appl
41	113.8	11.1	15056	34 US-60-029-960-31	Sequence 31, Appl
42	112.6	10.9	1101	1 PCT-US98-25010-1	Sequence 1, Appl 1
43	112.6	10.9	1101	13 US-08-979-616-1	Sequence 1, Appl 1
44	112.6	10.9	3417	1 PCT-US98-24955-7	Sequence 7, Appl 1
45	112.6	10.9	3417	1 PCT-US98-24955-7	Sequence 7, Appl 1

## ALIGNMENTS

```
RESULT 1
PCT-US98-24857-1
: Sequence 1, Application PC/TUS9824857A
: GENERAL INFORMATION:
: APPLICANT: SmithKline Beecham Corporation
: TITLE OF INVENTION: Novel ribg
: FILE REFERENCE: P50444-9
: CURRENT APPLICATION NUMBER: PCT/US98/24857A
: CURRENT FILING DATE: 1998-11-23
: EARLIER APPLICATION NUMBER: 08/978,456
: EARLIER FILING DATE: 1997-11-25
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: FASTSEQ for Windows Version 3.0
: SEQ ID NO 1
: LENGTH: 1029
: TYPE: DNA
: ORGANISM: Staphylococcus aureus
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)...(1026)
: PCT-US98-24857-1
```

```
Query Match 100.0%; Score 1029; DB 1; Length 1029;
Best Local Similarity 100.0%; Pred. No. 6.ee-224;
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 ATGATTTATGCGATTCACACTTTCGAATATGTACAAAGTCGAACAGGTGTTATCCACC 60
DB 1 atgattatgcatcacttgcgaatattgtacaaagtcgaacaggtgttatccacc 60
OY 61 GTTGGCGCTGTGTGTTAATGAAGTAGAGATTGTTGATTGTCACACTTGGAGAAA 120
DB 61 gttggcgctgtgtgttaatgaagtagagattgttgattgtcacacttggagaaa 120
OY 121 GGTGCAAGCATCGCAGGCTTCACGACTTGTATGCGCAACAACAATGCTGAAGTGG 180
DB 121 ggtgcaagcatcgcgaggttcacgacttgtatgycacaacaatgctgaagtgg 180
OY 181 ACGATTTATGATGCTTACGCTGATGTCATTTGTTCAACCCACCGTGTATAC 240
DB 181 acgatttatgacttaccgctgattgtcatttgttcaacccacccgtgtatac 240
OY 241 AAAATTTGATTTAGATAGCAAAAGTAGTATAGCAACAACAAGACATTCGTTAGAC 300
DB 241 aaattttgatattagatagcaaaagtagtattagcaacaacaagacatttcgttagac 300
OY 301 ACACATGATGATGACGCTTACGGCTCACGGTATTGACGTTGAATGCGTTGATGATA 360
DB 301 acacatgattgatgacgcttaccgctcacggatttgacgttgaatgctgtgatgata 360
OY 361 CGGCGATCACAATTATACCAAGCTTTTAAAGCAAAAGCAACCAATGCGACAAT 420
DB 361 cggcgatcacaaattataccagctttttaaagcaaaagcaaaccaatgcgacaat 420
OY 421 ACAGTGAAGATATGCAAGTTTAGATGTTAAACAAGCAAGATGATGATGCAAAAGTCAA 480
DB 421 acagtgaagatattgcaagtttagatgtttaaacaagcaagatgattgataaagtc 480
OY 481 TGGATTACTAACAAAGAGGTTAACACAGATGCTATTAAGTTAAGACATGACACGACA 540
DB 481 tggattactaacaaagaggttaacaagatgctataagtttaagacatgcacagaca 540
OY 541 GTGTTAAGTGAAGAGGTTAAGTGAATAGATGATGATGATGATGATGATGATGATGAT 600
DB 541 gtgttaagtgaagaggttaagtgaaatagatgattgattgattgattgattgattgatt 600
OY 601 GATGCAAAAAACCCATATAAAGTAAATATATGTTAGTGGGATATTCATTTATATG 660
DB 601 gatgcaaaaaacccatataaagtaaatattgttctaagttcgggaatatattcaatc 660
```

```
OY 661 CAAATTTATCAAGATGATCAACACCAATTGGATATATCTGAAATTCAAATTTACA 720
DB 661 caatttatcaagatgattcaacacaaattggatattattctgaaattccaattaca 720
OY 721 AGCATTCAAACATATGAAATTTATCTGAGAGCTGTTGATTTAAACAAATCTCT 780
DB 721 agcatccaacatattgaaatttatctgagagctgttgattttaaacaattctc 780
OY 781 CACAATTTATATAAAGAGAGAGTGAAGCTTGTGATGAGAGAGAGTGAAGAGTGA 840
DB 781 cacaatttatataaagagagagtgaaagcttgtgattgagagagagtgaaagag 840
OY 841 TCAGATTCGATCTATTAATATAGATGATGATGATGATGATGATGATGATGATGAT 900
DB 841 tcagatttcgatctattatataagatgattgattgattgattgattgattgattgatt 900
OY 901 GCGGATCTGGAATTTATCAATTTATCAAAATGATGATGATGATGATGATGATGAT 960
DB 901 gcggatctggaaatttatcaattttatcaaaatgatgatgatgatgatgatgatgat 960
OY 961 AACCAATTTGAATTTGATTCATTCGAGTTATTAATCAAAATGTTAATTAATTA 1020
DB 961 aaccaatttgaatttgattcattccgagttattataatcaaaatgttaatttaact 1020
OY 1021 AAGAGTGA 1029
DB 1021 aagaagtga 1029
```

```
RESULT 2
US-09-369-700-1
: Sequence 1, Application US/09369700
: GENERAL INFORMATION:
: APPLICANT: Palmer, Leslie M.
: TITLE OF INVENTION: Novel ribg
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESS: Dechert Price & Rhoads
: STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
: CITY: Philadelphia
: STATE: PA
: COUNTRY: US
: ZIP: 19103
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/369,700
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/978,456
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Dickinson, Todd Q
: REGISTRATION NUMBER: 28,354
: REFERENCE/DOCKET NUMBER: P50444-9
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 215-994-2252
: TELEFAX: 215-994-2222
: TELEX:
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1029 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: US-09-369-700-1
```

Query Match 100.0%; Score 1029; DB 17; Length 1029;  
Best Local Similarity 100.0%; Pred. No. 6,6e-224;  
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGATATATGATTCACCTTGCATTAATGTGTACAGTCAACAGGTGTTAATCCACC 60  
DB 1 ATGATATATGATTCACCTTGCATTAATGTGTACAGTCAACAGGTGTTAATCCACC 60  
QY 61 GTTGCGCGCTGTGTAGTAAATGAAGTAGATTTGTTGTTGTCACACTTGAGAAA 120  
DB 61 GTTGCGCGCTGTGTAGTAAATGAAGTAGATTTGTTGTTGTCACACTTGAGAAA 120  
QY 121 GGTGACACAGATGCGGAGGTTCAAGCTTGATATGCGACACAAATGCTGAAGTGG 180  
DB 121 GGTGACACAGATGCGGAGGTTCAAGCTTGATATGCGACACAAATGCTGAAGTGG 180  
QY 181 ACGATTTATTTACGTAGAGCCATGTAGTATTTGTTGTCACACACCTGTGTAC 240  
DB 181 ACGATTTATTTACGTAGAGCCATGTAGTATTTGTTGTCACACACCTGTGTAC 240  
QY 241 AAAATTTATTTAGTATAGTACAAAGTAGTATACGACAAAGCAATTCGTTAGAC 300  
DB 241 AAAATTTATTTAGTATAGTACAAAGTAGTATACGACAAAGCAATTCGTTAGAC 300  
QY 301 ACACATGCTGATGACAGCTTACGGGCTCACGGTATGAGTTGAATCGTTGATGAA 360  
DB 301 ACACATGCTGATGAGAGCTTACGGGCTCACGGTATGAGTTGAATCGTTGATGAA 360  
QY 361 CGGCGATCACAATTTATCCAAAGCTTTTAAAGCAAAAGCAAGTCCCAAT 420  
DB 361 CGGCGATCACAATTTATCCAAAGCTTTTAAAGCAAAAGCAAGTCCCAAT 420  
QY 421 ACAATGAAATATCTGCAATTTGATGTAACACAGCATGATATGACAAAGTCAA 480  
DB 421 ACAATGAAATATCTGCAATTTGATGTAACACAGCATGATATGACAAAGTCAA 480  
QY 481 TGGATTTCTAACAAGAGGTTAAACAGATGCTATTAAGTTAAGCATGACAGACGA 540  
DB 481 TGGATTTCTAACAAGAGGTTAAACAGATGCTATTAAGTTAAGCATGACAGACGA 540  
QY 541 GTGTAACTGAGAGCTGACATGATGATGATGATGATGATGATGATGATGATGAT 600  
DB 541 GTGTAACTGAGAGCTGACATGATGATGATGATGATGATGATGATGATGATGAT 600  
QY 601 GATGAGAAAAACCTTATAAAGTATATGCTAGCTGGGATTTTCAATTTAATCAG 660  
DB 601 GATGAGAAAAACCTTATAAAGTATATGCTAGCTGGGATTTTCAATTTAATCAG 660  
QY 661 CAAATTTATCAGATGATCAACACCAATTTGATATATCTGAAATTCGCAATTTACA 720  
DB 661 CAAATTTATCAGATGATCAACACCAATTTGATATATCTGAAATTCGCAATTTACA 720  
QY 721 AGCAATTAACACATATTTGAATTTTACTTGAAGTCTGTGATTAACACATTTCT 780  
DB 721 AGCAATTAACACATATTTGAATTTTACTTGAAGTCTGTGATTAACACATTTCT 780  
QY 781 CACATTTATTAATAAGAGAGTGTGAAGTCTGATGAGGAGGTCACACACTACT 840  
DB 781 CACATTTATTAATAAGAGAGTGTGAAGTCTGATGAGGAGGTCACACACTACT 840  
QY 841 TCAGATTTTCCATCTATTAATAGATGAATTAATTCCTATTAATGCCCCGAAATTA 900  
DB 841 TCAGATTTTCCATCTATTAATAGATGAATTAATTCCTATTAATGCCCCGAAATTA 900  
QY 901 GGGGATCTGGAATTTTCAATTTTATCAAAATGATGATGATGATGATGATGATGAT 960  
DB 901 GGGGATCTGGAATTTTCAATTTTATCAAAATGATGATGATGATGATGATGATGAT 960  
QY 961 AACCAATTTGAATTTGTTCACTTCGAGTATTAATCAAAATGTTAATTAATTAAT 1020  
DB 961 AACCAATTTGAATTTGTTCACTTCGAGTATTAATCAAAATGTTAATTAATTAAT 1020  
QY 1021 AACCAATTTGA 1029

DB 1021 AACCAATTTGA 1029

RESULT 3  
US-09-376-633-1  
; Sequence 1, Application US/09376633  
; GENERAL INFORMATION:  
; APPLICANT: Palmer, Leslie M.  
; TITLE OF INVENTION: Novel rldg  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dechert Price & Rhoads  
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: US  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/376,633  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/978,456  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dickinson, Todd O  
; REGISTRATION NUMBER: 28,354  
; REFERENCE/DOCKET NUMBER: P50444-9  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-994-2252  
; TELEFAX: 215-994-2222  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1029 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; US-09-376-633-1

Query Match 100.0%; Score 1029; DB 17; Length 1029;  
Best Local Similarity 100.0%; Pred. No. 6,6e-224;  
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGATATATGATTCACCTTGCATTAATGTGTACAGTCAACAGGTGTTAATCCACC 60  
DB 1 ATGATATATGATTCACCTTGCATTAATGTGTACAGTCAACAGGTGTTAATCCACC 60  
QY 61 GTTGCGCGCTGTGTAGTAAATGAAGTAGATTTGTTGTTGTCACACTTGAGAAA 120  
DB 61 GTTGCGCGCTGTGTAGTAAATGAAGTAGATTTGTTGTTGTCACACTTGAGAAA 120  
QY 121 GGTGACACAGATGCGGAGGTTCAAGCTTGATATGCGACACAAATGCTGAAGTGG 180  
DB 121 GGTGACACAGATGCGGAGGTTCAAGCTTGATATGCGACACAAATGCTGAAGTGG 180  
QY 181 ACGATTTATTTACGTAGAGCCATGTAGTATTTGTTGTCACACACCTGTGTAC 240  
DB 181 ACGATTTATTTACGTAGAGCCATGTAGTATTTGTTGTCACACACCTGTGTAC 240  
QY 241 AAAATTTATTTAGTATAGTACAAAGTAGTATACGACAAAGCAATTCGTTAGAC 300  
DB 241 AAAATTTATTTAGTATAGTACAAAGTAGTATACGACAAAGCAATTCGTTAGAC 300  
QY 301 ACACATGCTGATGACAGCTTACGGGCTCACGGTATGAGTTGAATCGTTGATGAA 360  
DB 301 ACACATGCTGATGAGAGCTTACGGGCTCACGGTATGAGTTGAATCGTTGATGAA 360

```

Db 301 ACACATGATGAGACGTTACGGCTCACGGTATTGAGCTGTAATGCGTTGATGTGA 360
Qy 361 CGGGCATCACAATTATVACCAAGACTTTTAAAGCAAAAGCAAGCACTGCCCAATT 420
Db 361 CGGGCATCACAATTATVACCAAGACTTTTAAAGCAAAAGCAAGCACTGCCCAATT 420
Qy 421 ACAGTGAAGATCTGCAAGTTAGTGTAAACAAAGCAATGATTAAGACAAAGTCA 480
Db 421 ACAGTGAAGATCTGCAAGTTAGTGTAAACAAAGCAATGATTAAGACAAAGTCA 480
Qy 481 TGGATTACTAACAAAGAGTTAAACAGATGCTATAGTTAAGATGACGACGACGA 540
Db 481 TGGATTACTAACAAAGAGTTAAACAGATGCTATAGTTAAGATGACGACGACGA 540
Qy 541 GTGTTAACTGGAAGACGTACAGTTGATTAAGATGATCCAAATATACACGATTTCA 600
Db 541 GTGTTAACTGGAAGACGTACAGTTGATTAAGATGATCCAAATATACACGATTTCA 600
Qy 601 GATGGAAGAAACCCCTATTAAGATATTTGCTAAGCTGGGAAATTTCAATTACG 660
Db 601 GATGGAAGAAACCCCTATTAAGATATTTGCTAAGCTGGGAAATTTCAATTACG 660
Qy 661 CAAATTTCAAGATGATCAACACCAATTTGATATACGAAATCCAAATTTACA 720
Db 661 CAAATTTCAAGATGATCAACACCAATTTGATATACGAAATCCAAATTTACA 720
Qy 721 ACCAATCAACAACATATTTGAATTTATTTACTGAACTCTTGATTTAAACAACATTTCT 780
Db 721 ACCAATCAACAACATATTTGAATTTATTTACTGAACTCTTGATTTAAACAACATTTCT 780
Qy 781 CCAATTTATTAAGAGAGAGTGTGACTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db 781 CCAATTTATTAAGAGAGAGTGTGACTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Qy 841 TCGAATTTCCATCTATTTATATAGATGATTTATTTCTATTTATCCCGCAATTTAAT 900
Db 841 TCGAATTTCCATCTATTTATATAGATGATTTATTTCTATTTATCCCGCAATTTAAT 900
Qy 901 GCGGATCTGGAATTTATCAATTTTATCAACAACATGATGATGATGATGATGATG 960
Db 901 GCGGATCTGGAATTTATCAATTTTATCAACAACATGATGATGATGATGATGATG 960
Qy 961 AACCAATTTGAATGTTGATCCGAGTATTAATCAAAATGTTAAATTAATTAATTA 1020
Db 961 AACCAATTTGAATGTTGATCCGAGTATTAATCAAAATGTTAAATTAATTAATTA 1020
Qy 1021 AAGAGTGA 1029
Db 1021 AAGAGTGA 1029

```

```

RESULT 4
PCT-US98-24857-7
: Sequence 7, Application PC/TUS9824857A
: GENERAL INFORMATION:
: APPLICANT: SmithKline Beecham Corporation
: FILE OF INVENTION: Novel r1d6
: CURRENT APPLICATION NUMBER: PCT/US98/24857A
: CURRENT FILING DATE: 1998-11-23
: EARLIER APPLICATION NUMBER: 08/978,456
: EARLIER FILING DATE: 1997-11-25
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 7
: LENGTH: 3336
: TYPE: DNA
: ORGANISM: Staphylococcus aureus
PCT-US98-24857-7

```

```

Query Match 100.0%; Score 1029; DB 1; Length 3336;
Best Local Similarity 100.0%; Pred. No. 9.2e-224;

```

```

Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGATTATGCGATTTCACCTGCAATATGTACAAAGTCAAGAGTGTATATCCACC 60
Db 1 atgattatgcgatttcaacttgcgaataatgatacaaggtcaacagggttataccacc 60
Qy 61 GTTGGCGTGTGTAGTTAATGAAGGTAGATTTGTTGATTTGTGCACATTGAGAAA 120
Db 61 gttggcggtgtgtagttaataagaaggttagatttggatgttgcacattgagaaaa 120
Qy 121 GGTGAACACATGCGGAGGTTCACGACCTGATATGGCAACAAATGCTAAGTGGC 180
Db 121 ggtgaacacatgcgaggttcacgacctgatatggcacaatgctaaagtggc 180
Qy 181 ACGATTTATATTAAGTAAAGCCATGATCTATTTGTTTAAACACCACTGTGTAA 240
Db 181 acgatttatataagtaagccatgattctttttaaagcaaacaccctgtgtaac 240
Qy 241 AAAATTTGATTTAGATAGCAAAAAGTATACGCAAAAAGACAAATTTGTTAGC 300
Db 241 aaatattgatttgaagtagaagcaaaagtagatacgcaaaaagacattcgttagc 300
Qy 301 ACACATGATGAGAGAGAGTACGGGCTCACGGTATTGAGTGTGAATGCTGTGATGA 360
Db 301 acacatgagtagagagagtagcggctcacggatttgagtgatgagtagatga 360
Qy 361 CGGGCATCACAATTATVACCAAGACTTTTAAAGCAAAAGCAAGCACTGCCCAAA 420
Db 361 cgggcatacacaattatvaccagactttttaaagcaaaagcaagcaactgccaat 420
Qy 421 ACAGTGAAGATCTGCAAGTTAGTGTAAACAAAGCAATGATTAAGACAAAGTCA 480
Db 421 acagtgaagatctgcagagtttagatgtaaacaagcgaatgataatgaaagttca 480
Qy 481 TGGATTACTAACAAAGAGTGTAAACAGATGCTATTAAGTAAAGATGACGACGACGA 540
Db 481 tggattactaacaaagaggttaaacagatgctataaagtaagaatcgacga 540
Qy 541 GTGTTAACTGGAAGACGTACAGTTGATTAAGATGATCCAAATATACACGATTTCA 600
Db 541 gtttaactggaagagcgtacagttgaattagatgccaatatactacagattca 600
Qy 601 GATGGAAGAAACCCCTATTAAGATATTTGCTAAGCTGGGAAATTTCAATTACG 660
Db 601 gatggaagaaacccctatataagatatttgctaaagctgggaatttcaattacg 660
Qy 661 CAAATTTCAAGATGATCAACACCAATTTGATTAAGTATTAAGTAAATTTAACA 720
Db 661 caaatttcaagatgatacacaaccaatttgatatacaccgaatlttaaca 720
Qy 721 ACCAATCAACAACATATTTGAATTTATTTACTGAACTCTTGATTTAAACAATTTCT 780
Db 721 accaatcaacaacatatttgaatttatttactggaactcttgatlttaacaacttct 780
Qy 781 CCAATTTATTAAGAGAGAGTGTGACCTTTGCTAAGTGAAGAGAGTGCACACCTACT 840
Db 781 ccaatttataataaagagagtgtaacttgcctagtgagagagtgcaactact 840
Qy 841 TCGAATTTCCATCTATTTATATAGATGATTTATTTCTATTTATCCCGCAATTTAAT 900
Db 841 tcgaatttccatctatattatataagatgattatcttctatcagcccgaaatlaa 900
Qy 901 GCGGATCTGGAATTTATCAATTTTATCAACAACATGATGATGATGATGATGATG 960
Db 901 gcggatctggaatttatcaattttatcaacaacatgatgatgatgatgagatgag 960
Qy 961 AACCAATTTGAATGTTGATCCGAGTATTAATCAAAATGTTAAATTAATTAATTA 1020
Db 961 aaccaattggaattgattcaccgagtattataaatacaaatgtaaatttaactta 1020
Qy 1021 AAGAGTGA 1029
Db 1021 aagaagtga 1029

```





Qy	361	CGGGCATCACAAATTAATCCAAAGACTTTTAAAGCAAAAGCAACCAACTGGCAAAATT	420
Dp	361	cgggcatcacaaattataccaagaacttlttaaagcaaaagcaactgcgcaaat	420
Qy	421	ACAGTAAGATATCTGCAGTTAGATGGTAACAAGCAATGATATGACAAAGTAA	480
Dp	421	acaatgaaatatactgcaagcttagatggtaaacaagcgaaatgataatgacaagtc	480
Qy	481	TGCATTAAGTAACAAGAGGTTTAAACAAGATGTCCTAATTAATTAAAGCATCGACAGACCA	540
Dp	481	tgcattatacaaaagaggttaacaagaatgcttaagaatgaagacacgaacgaagca	540
Qy	541	GTGTTAACTGGAAGACGTACAGTTGAATTAGATGATCCAAATPATACACGATTTCAA	600
Dp	541	gtgttaactggaagacgttacagctggaattagatgatccacaataatactaacgta	600
Qy	601	GATGGAAAAAACCCATATAAAGTAATATTGTCTAAGTCTGGGAAATATTCATTTAATAG	660
Dp	601	gataggaaaaaacccataaaagtaaatatgctctaaagctcggaataatcatat	660
Qy	661	CAAAATTAATACATGATCAATCAACCAATTTGGATTAATACGAAATATCCAAATTAA	720
Dp	661	caaatatataagaatgaatcaacaacccaatttggatataatacggaaaacccaatt	720
Qy	721	AGCAATTAACACATATTTGAATTATTACTTGAAGTCTTGGAATTAAACAAATTTCT	780
Dp	721	agcaataaacaacataatgaataatcttaacttgaagctcttggatttaacaaca	780
Qy	781	CACAAATTAATATAAGAGAGAGATTTGGACCTTGGCTGTGAGAGCGATCCAAACCTACT	840
Dp	781	cacaattataataaagagagcttgaagaactctgcacgacgagcgagcccaac	840
Qy	841	TCAGATATTCACATCTAATATATATAGATGAATTAATTTCTATTAAGCCCGAAATTAAT	900
Dp	841	tcgaatattccacatctatataataagatgaattatctctatataagcccgaa	900
Qy	901	GGCGATCTCGAAATTAATCAATTTTTCAAACAATGATGATTAAGATACCGATCG	960
Dp	901	ggcgatctcgaaatatatacaattctcaacaagaatgagatctggaatccgaag	960
Qy	961	AACCAATTTGAATTTGTTCAATTCGAGTTATTAAATCAAAATGTTAAATTAAC	1020
Dp	961	aaccaattgaaatgctacccgagctattaaatcaaaatgtaaat	1020
Qy	1021	AAGAAGTGA 1029	
Dp	1021	aagaagctga 1029	

```

RESULT 7
PCT-US98-25096-5
: Sequence 5: Application PC/TUS9825096
: GENERAL INFORMATION:
: APPLICANT: SmithKline Beecham Corporation
: TITLE OF INVENTION: r1b3
: FILE REFERENCE: GM10123
: CURRENT APPLICATION NUMBER: PCT/US98/25096
: CURRENT FILING DATE: 1998-11-24
: EARLIER APPLICATION NUMBER: 08/977,553
: EARLIER FILING DATE: 1997-11-25
: NUMBER OF SEQ ID NOS: 5
: SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 5
LENGTH: 3536
:
: TYPE: DNA
: ORGANISM: Staphylococcus aureus
PCT-US98-25096-5

```

Query Match	100.0%	Score 1029;	DB 1;	Length 3336;
Best Local Similarity	100.0%	Pred. No. 9.2e-224;		
Matches 1029;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	1	ATGATTTATGCGAATTCACCTGCGAAATATGATGACAAAGGTCCAAACAGGTTAATCCACC	60
Db	1	atgattttatgcgatccaacttgcgaatacgttaacaaggctcaacaaggctgtaataccaacc	60
Qy	61	GTTGGCCCTTTTGTAGTTAATGAAGGTACGATTTGTTGGTATTTGGTGCACCTTGAGAAA	120
Db	61	gttgcccttctgtagtttaatagaaggtacgatttggatttgggtgcaccttgaganaa	120
Qy	121	GCTGACAAAGCATCCGAGAGTTCAAGCAGCTTGATATGCGACAAACAAATGCTGAAGGTGCG	180
Db	121	gctgacaagaatccggaaggttcaaggaacttgaatgacgaacaacaatgtcgaaggtgcg	180
Qy	181	ACGATTTATTTATACGTTAGAGCCATGTATGCAATTTTGGTTCAACACCACTCGTGTATAC	240
Db	181	acgatttataatcgtltagtagcgaatgtatgtcaatttgglttcaacaccacctgtgttaac	240
Qy	241	AAATATTTGATTTGTAGATGACAAAAGTAGTATACGCACAAAAGCAATTCGTTAGAC	300
Db	241	aaaattatgtattgtaagaatgacaaaagtagatatacgcacaaaagaacattcgttagac	300
Qy	301	ACACATGCTATGAGCGTTACGGGCTCACGGTATTTGGATTTGGATTCGTTGATGATGA	360
Db	301	acacatgctatgagcgttaccggctcacggatatttggatttggatttcggtgatgatga	360
Qy	361	CGGGCATCACAAATTTATACACAGACTTTTAAAGCAAAAGCAAAAGCAATTCGCCAAATT	420
Db	361	cgggcatacaaatattacacagactttttaaagcaaaaagcaaacatctgccacaatt	420
Qy	421	ACAGTGAATATATCTGCAGAGTTTAATGTTGTTAAACAAGCAATGATATGACAAAGTCGA	480
Db	421	acagtgaatattctgcagagtttaattggtttaaacaagcaatgatattgacaaagtcca	480
Qy	481	TGCATTTACTTACCAAAAGAGTTTAAACAAGATGCTTATGTTAATGAACATCTCACAGCGCA	540
Db	481	tgcattttacttaacaaagggtttaaacaagatgtcctaagaatgaacatlgacaacgcgca	540
Qy	541	GTGTTAACTGGAAGAGCTACAGTTGAATTTAGATGATCCACAATTTACTPACAGTATTCAA	600
Db	541	gtgttaacttggaaagactacagttgaattagatgtatgccacaatatactacaagttacca	600
Qy	601	GATGCAAAAACCCCTTAAAGAAATATTTGCTGATGCTGGGATTTTCATTTTAATCAG	660
Db	601	gatgcaaaaaccccttaaaagaaatatttgcctgacgtccggaaatctcatcttcaacag	660
Qy	661	CAAAATTTATCAAGATCAATCAACACCAATTTGGATATATACTGAAAATTCCAAAATTTAACA	720
Db	661	caaatattatcaagaatgaatccaacccaatttggatatactatgtaaatccaatttaaca	720
Qy	721	AGCAATTAACACATATTTGAAATTTATTTACTTGAAGTTTGTGATTTAACACAATTCCT	780
Db	721	agcaattaacaacatatgaaattatctacttgaaggtctgtgattttaaacaacattctt	780
Qy	781	CACAAATTTATTAAGAGAGATTTGAACTTTGGATGTCGAGCGAGAGTCCACACACTACT	840
Db	781	caaatattataataaagaagagtttgaacttctcagtcgagcggaaggtccaaccactact	840
Qy	841	TCAGAAATTCCTCACTTATTTATATAGATGAATTTATTTCTATTTATGGCCGGAATTAATT	900
Db	841	tcgaattatcccatctattatataagatgaattatctctatacttaccgccaatttaatt	900
Qy	901	GGGGGATCTGGAATTTTCAATTTTATCAACAAATGATGTTGGATTTGATACCAAGATGGG	960
Db	901	ggcggaactcggaaattatcaattatcttaacaacaagtgtgtgatgataccagatgcg	960
Qy	961	AACCAATTTGAAATTTGTCATTTCCGAGTTTAAATCAAAATGTTAAATTAATTAATTTACGA	1020
Db	961	aaccaatattgaatattgttcatctcgagttatataacaagaatgttaaatctaactttacga	1020
Qy	1021	AAGAAGTGA 1029	
Db	1021	aagaagtgga 1029	

```
RESULT 8
US-08-977-553-5
: Sequence 5, Application US/08977553
: GENERAL INFORMATION:
: APPLICANT: Palmer, Leslie M.
: TITLE OF INVENTION: Novel ribb
: NUMBER OF SEQUENCES: 5
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dechart Price & Rhoads
: STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
: City: Philadelphia
: STATE: PA
: COUNTRY: US
: ZIP: 19103
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/977,553
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Dickinson, Todd O
: REGISTRATION NUMBER: 28,354
: REFERENCE/DOCKET NUMBER: GM10123
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 215-994-2252
: TELEFAX: 215-994-2222
: TELEX:
: INFORMATION FOR SFO ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3336 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: US-08-977-553-5

Query Match 100.0%; Score 1029; DB 13; Length 3336;
Best Local Similarity 100.0%; Pred. No. 9, 2e-224;
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

Db 361 CGGGATCACCATTATACCAAGACTTTTAAAGCAAAAGCAACGTCACCAAAATT 420
Qy 421 ACAGTGAAGATATCTGCAAGTTTAGATGTAACAGCAATGATTAATGCAAGTCAA 480
Db 421 ACAGTGAAGATATCTGCAAGTTTAGATGTAACAGCAATGATTAATGCAAGTCAA 480
Qy 481 TGGATTACTAACAAGAGTTAAACAGATGCTATAAGTTAAGACATCGACAGCGCA 540
Db 481 TGGATTACTAACAAGAGTTAAACAGATGCTATAAGTTAAGACATCGACAGCGCA 540
Qy 541 GTGTTAACTGGAAGACGTACAGTTGAATATGATATCCAAATATACACGATTCAA 600
Db 541 GTGTTAACTGGAAGACGTACAGTTGAATATGATATCCAAATATACACGATTCAA 600
Qy 601 GATGAAAAAACCCCTATTAAGTATATGCTAAGCTGGGAAATATCATTTAAATCAG 660
Db 601 GATGAAAAAACCCCTATTAAGTATATGCTAAGCTGGGAAATATCATTTAAATCAG 660
Qy 661 CAAATTTATCAAGATGAATCAACACCAATTTGATATATACTGAAATCCAAATTTACA 720
Db 661 CAAATTTATCAAGATGAATCAACACCAATTTGATATATACTGAAATCCAAATTTACA 720
Qy 721 AGCAATCAACACATATTTGAATTTACTTGAAGTCTTGATTTAACACAAATCTT 780
Db 721 AGCAATCAACACATATTTGAATTTACTTGAAGTCTTGATTTAACACAAATCTT 780
Qy 781 CACAATTTATTAAGAGAGGATGGAACCTTGTAGTCGAGGAGGACCAACCACTACT 840
Db 781 CACAATTTATTAAGAGAGGATGGAACCTTGTAGTCGAGGAGGACCAACCACTACT 840
Qy 841 TCAGAAATTCCTCATCTATTAATAGATGAATTTATCTATTAATGCCCCGAAATTAAT 900
Db 841 TCAGAAATTCCTCATCTATTAATAGATGAATTTATCTATTAATGCCCCGAAATTAAT 900
Qy 901 GCGGATCTGGAATTTATCAATTTTATCAAAACAAATGATGATTAATCCAGATGCG 960
Db 901 GCGGATCTGGAATTTATCAATTTTATCAAAACAAATGATGATTAATCCAGATGCG 960
Qy 961 AACCAATTTGAATTTGATTCGAGTATTAATCAAAATGTTAAATTAATTAATTCGA 1020
Db 961 AACCAATTTGAATTTGATTCGAGTATTAATCAAAATGTTAAATTAATTAATTCGA 1020
Qy 1021 AAGAAGTGA 1029
Db 1021 AAGAAGTGA 1029

RESULT 9
US-08-977-866-7
: Sequence 7, Application US/08977866
: GENERAL INFORMATION:
: APPLICANT: Palmer, Leslie M.
: TITLE OF INVENTION: Novel ribb
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dechart Price & Rhoads
: STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
: City: Philadelphia
: STATE: PA
: COUNTRY: US
: ZIP: 19103
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/977,866
: FILING DATE:
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
```

```

: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Dickinson, Todd O
: REGISTRATION NUMBER: 28, 354
: REFERENCE/DOCKET NUMBER: P50444-8
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 215-994-2252
: TELEFAX: 215-994-2222
:
: TELETYPE:
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3336 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
:
: US-08-977-866-7

```

```

Query Match      100.0%; Score 1029; DB 13; Length 3336;
Best Local Similarity 100.0%; Pred. No. 9.2e-224;
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGATTATGCGATTCACATTCGCAATATGTGTACAGGTCAACAGGTGTTAATCCACCC 60
DB 1 ATGGATTATGCGATTCACATTCGCAATATGTGTACAGGTCAACAGGTGTTAATCCACCC 60
QY 61 GTTGGCGCTGTTGTAGTTAATGAAGGTAGCATTTGCTATTGCTGACACTTGAGAAAA 120
DB 61 GTTGGCGCTGTTGTAGTTAATGAAGGTAGCATTTGCTATTGCTGACACTTGAGAAAA 120
QY 121 GGTGACACATGCGGAGGTTCACAGCACTTGATATGCGCAACAAATGCTGAAGTGGG 180
DB 121 GGTGACACATGCGGAGGTTCACAGCACTTGATATGCGCAACAAATGCTGAAGTGGG 180
QY 181 AGCATTTATATTACGCTTAGAGCATGTAGTCATTTGTTCAACACACCCCTGTGTAA 240
DB 181 AGCATTTATATTACGCTTAGAGCATGTAGTCATTTGTTCAACACACCCCTGTGTAA 240
QY 241 AAAATTATTGATTGTAAGATAGCAAAAGTAGTATACGCAACAAAGACAATTCGTTAGAC 300
DB 241 AAAATTATTGATTGTAAGATAGCAAAAGTAGTATACGCAACAAAGACAATTCGTTAGAC 300
QY 301 AACATGTTGATGAGACGTTAGCGGCTCAGCGTATTGAGGTGAATGCTGTATGAA 360
DB 301 AACATGTTGATGAGACGTTAGCGGCTCAGCGTATTGAGGTGAATGCTGTATGAA 360
QY 361 CGGCGATCACAATTTATACCAAGACTTTTAAAGCAAAAGCAAGCGACCAAAAT 420
DB 361 CGGCGATCACAATTTATACCAAGACTTTTAAAGCAAAAGCAAGCGACCAAAAT 420
QY 421 ACAGTGAAGTAGTTCGCAAGTTAGATGTAACAAGCAAGTATGATGACAAAGTCAA 480
DB 421 ACAGTGAAGTAGTTCGCAAGTTAGATGTAACAAGCAAGTATGATGACAAAGTCAA 480
QY 481 TGGATTACTAACAAGAGGTAAACAAGTGTCTAATAGTAAAGCATGACACGCGCA 540
DB 481 TGGATTACTAACAAGAGGTAAACAAGTGTCTAATAGTAAAGCATGACACGCGCA 540
QY 541 GTGTTAACTGAAGACGTACAGTGAATAGATGATCCAAATATCTACAGTATTCAA 600
DB 541 GTGTTAACTGAAGACGTACAGTGAATAGATGATCCAAATATCTACAGTATTCAA 600
QY 601 GATGAAAAAACCCCTATAAAGTATATGCTAAGCTGCGGAATTTGATTTAATCAG 660
DB 601 GATGAAAAAACCCCTATAAAGTATATGCTAAGCTGCGGAATTTGATTTAATCAG 660
QY 661 CAAATTTATCAAGATGATCAACACCAATTTGATATATGAAATTCCTAATTTAACA 720
DB 661 CAAATTTATCAAGATGATCAACACCAATTTGATATATGAAATTCCTAATTTAACA 720
QY 721 ACCAATCAACACATTTGAAATTTACTGGAAGTCTTGATTTAACAACAATTCCT 780
DB 721 ACCAATCAACACATTTGAAATTTACTGGAAGTCTTGATTTAACAACAATTCCT 780

```

```

QY 781 CACAATTTATATAAAGAGAGTTGGACCTTGTAGTGCAGGACGTCACCACTACT 840
DB 781 CACAATTTATATAAAGAGAGTTGGACCTTGTAGTGCAGGACGTCACCACTACT 840
QY 841 TCAGAAATTCATCTATTATATAGATGAATTTATTTCTTATTATGCCCGAAATTAAT 900
DB 841 TCAGAAATTCATCTATTATATAGATGAATTTATTTCTTATTATGCCCGAAATTAAT 900
QY 901 GCGGATCTGCAAAATATATCAACAAATGATGTGATGAGATACCAGATGCG 960
DB 901 GCGGATCTGCAAAATATATCAAAATGATGTGATGAGATACCAGATGCG 960
QY 961 AACCAATTTGAAATTTGTTATTCGAGTATTAATCAAAATGTTAAATTAATTTAGCA 1020
DB 961 AACCAATTTGAAATTTGTTATTCGAGTATTAATCAAAATGTTAAATTAATTTAGCA 1020
QY 1021 AAGAACTGA 1029
DB 1021 AAGAACTGA 1029

```

```

RESULT 10
US-09-227-806-7
: Sequence 7, Application US/09227806
:
: GENERAL INFORMATION:
: APPLICANT: Palmer, Leslie M.
: APPLICANT: Pedon, Jason C.
: APPLICANT: Warren, Richard L.
: APPLICANT: Trainl, Christopher M.
: APPLICANT: Wang, Min
: APPLICANT: Jaworski, Deborah D.
: APPLICANT: Mooney, Jeffrey
: APPLICANT: Debouck, Christine
: APPLICANT: Zhong, Yixi
: APPLICANT: Black, Michael
: TITLE OF INVENTION: TIDA
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dechert, Price & Rhoads
: STREET: 4000 Bell Atlantic Tower, 1717 Arch Ste
: CITY: Philadelphia
: STATE: PA
: COUNTRY: USA
: ZIP: 19103-2793
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/227, 806
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/08/977, 554
: FILING DATE:
: APPLICATION NUMBER: PCT/US97/02318
: FILING DATE: 19-FEB-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Dickinson, O. Todd
: REGISTRATION NUMBER: 28, 354
: REFERENCE/DOCKET NUMBER: P50444-07
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 215/994-2252
: TELEFAX: 215/994-2222
:
: TELETYPE:
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3336 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear

```





|||||  
Db 661 CAAATTTATCAAGATGATCAACACCAATTTGATATATCTGAAAATCCAAATTTTACA 720  
QY 721 AGCAATCAAAACACATATTTGAATTTATTTACTTGAAGCTTGTGATTTTACAACATTTCT 780  
Db 721 AGCAATCAAAACACATATTTGAATTTATTTACTTGAAGCTTGTGATTTTACAACATTTCT 780  
QY 781 CACAATTTATTAATAAG 840  
Db 781 CACAATTTATTAATAAG 840  
QY 841 TCACAATTTCCATCTATTTATATAGATGATTTATTTCTTTATATGCCCCCAATTTAT 900  
Db 841 TCACAATTTCCATCTATTTATATAGATGATTTATTTCTTTATATGCCCCCAATTTAT 900  
QY 901 GCGGATCTGGAATTTATCAATTTTATCAACAATGATGATGATGATGATGATGATGATG 960  
Db 901 GCGGATCTGGAATTTATCAATTTTATCAACAATGATGATGATGATGATGATGATGATG 960  
QY 961 AACCAATTTGAATTTGTCATTCGAGTTATTAATCAAAATGTTAAATTAATTTACGTA 1020  
Db 961 AACCAATTTGAATTTGTCATTCGAGTTATTAATCAAAATGTTAAATTAATTTACGTA 1020  
QY 1021 AAGAAGTGA 1029  
Db 1021 AAGAAGTGA 1029

RESULT 13  
US-09-492-479-7

? Sequence 7, Application US/09492479  
? GENERAL INFORMATION:  
? APPLICANT: Palmer, Leslie M.  
? TITLE OF INVENTION: Novel ribb  
? NUMBER OF SEQUENCES: 7  
? CORRESPONDENCE ADDRESS:  
? ADDRESSEE: Dechert Price & Rhoads  
? STREET: 4000 Bell Atlantic tower, 1717 Arch Stre  
? CITY: Philadelphia  
? STATE: PA  
? COUNTRY: US  
? Zip: 19103  
? COMPUTER READABLE FORM:  
? MEDIUM TYPE: Diskette  
? COMPUTER: IBM Compatible  
? OPERATING SYSTEM: DOS  
? SOFTWARE: FastSeq for Windows Version 2.0  
? CURRENT APPLICATION DATA:  
? APPLICATION NUMBER: US/09/492,479  
? FILING DATE:  
? CLASSIFICATION:  
? PRIOR APPLICATION DATA:  
? APPLICATION NUMBER: 08/977,866  
? FILING DATE:  
? ATTORNEY/AGENT INFORMATION:  
? NAME: Dickinson, Todd Q  
? REGISTRATION NUMBER: 28,354  
? REFERENCE/DOCKET NUMBER: P50444-8  
? TELECOMMUNICATION INFORMATION:  
? TELEPHONE: 215-994-2252  
? TELEFAX: 215-994-2222  
? TELEX:  
? INFORMATION FOR SEQ ID NO: 7:  
? SEQUENCE CHARACTERISTICS:  
? LENGTH: 3336 base pairs  
? TYPE: nucleic acid  
? STRANDEDNESS: double  
? TOPOLOGY: linear  
? US-09-492-479-7

Query Match 100.0%; Score 1029; DB 18; Length 3336;  
Best Local Similarity 100.0%; Pred. No. 9, 2e-224;

Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGATTTATGCAATTCACATTCGAAATATGTGTACAGGTCAACAGGTGTTAATCCACC 60  
Db 1 ATGATTTATGCGATTCACATTCGAAATATGTGTACAGGTCAACAGGTGTTAATCCACC 60  
QY 61 GTTGGCGCTGTTGATTAATGAAGTAGATTTGTTGTTATTTGTCACACTTGAGAAA 120  
Db 61 GTTGGCGCTGTTGATTAATGAAGTAGATTTGTTGTTATTTGTCACACTTGAGAAA 120  
QY 121 GGTGACAGCATGCGGAGCTCAAGCACTTGATATGSCAACAAATGCTGAAGTCCG 180  
Db 121 GGTGACAGCATGCGGAGCTCAAGCACTTGATATGSCAACAAATGCTGAAGTCCG 180  
QY 181 AGCATTTATTTACGTTGAGCAGATGTATGTCATTTGTTCAACACCCCTGTTTAC 240  
Db 181 AGCATTTATTTACGTTGAGCAGATGTATGTCATTTGTTCAACACCCCTGTTTAC 240  
QY 241 AAAATTAATGATTTGTAAGATGCAAAAGTAGATATGCAACAAAGCAATTCGTTAGC 300  
Db 241 AAAATTAATGATTTGTAAGATGCAAAAGTAGATATGCAACAAAGCAATTCGTTAGC 300  
QY 301 ACACATGSGTAGAGACAGCTTACGGGCTCAGCGTATGAGTGTGAATGCGTTGATGATGA 360  
Db 301 ACACATGSGTAGAGACAGCTTACGGGCTCAGCGTATGAGTGTGAATGCGTTGATGATGA 360  
QY 361 CCGGATCACAATTAATACCAAGACTTTTAAAGCAAAAGCAACGCACTGCCAATTT 420  
Db 361 CCGGATCACAATTAATACCAAGACTTTTAAAGCAAAAGCAACGCACTGCCAATTT 420  
QY 421 ACAGTGAAGTATCTGCAAGTTGATGATGTTAAACAGGCAATGATATGACAAAGTCA 480  
Db 421 ACAGTGAAGTATCTGCAAGTTGATGATGTTAAACAGGCAATGATATGACAAAGTCA 480  
QY 481 TGGATTTACTACAAAGAGCTTAACAGATGTCATTAAGTTAAGCATCGACAGCCCA 540  
Db 481 TGGATTTACTACAAAGAGCTTAACAGATGTCATTAAGTTAAGCATCGACAGCCCA 540  
QY 541 GTGTTAACTGGAAGACGATGATGATGATGATGATGATGATGATGATGATGATGATG 600  
Db 541 GTGTTAACTGGAAGACGATGATGATGATGATGATGATGATGATGATGATGATGATG 600  
QY 601 GATGAAAAAACCCTATTAAGATATTTGCTAAGCTGGGAATATTCATTTAATCAG 660  
Db 601 GATGAAAAAACCCTATTAAGATATTTGCTAAGCTGGGAATATTCATTTAATCAG 660  
QY 661 CAAATTTATCAAGATGATCAACACCAATTTGATATTTACTGAAAAATTTTACA 720  
Db 661 CAAATTTATCAAGATGATCAACACCAATTTGATATTTACTGAAAAATTTTACA 720  
QY 721 AGCAATCAAAACACATATTTGAATTTATTTACTTGAAGCTTGTGATTTTACAACATTTCT 780  
Db 721 AGCAATCAAAACACATATTTGAATTTATTTACTTGAAGCTTGTGATTTTACAACATTTCT 780  
QY 781 CACAATTTATTAATAAG 840  
Db 781 CACAATTTATTAATAAG 840  
QY 841 TCACAATTTCCATCTATTTATATAGATGATTTATTTCTTTATATGCCCCCAATTTAT 900  
Db 841 TCACAATTTCCATCTATTTATATAGATGATTTATTTCTTTATATGCCCCCAATTTAT 900  
QY 901 GCGGATCTGGAATTTATCAATTTTATCAACAATGATGATGATGATGATGATGATGATG 960  
Db 901 GCGGATCTGGAATTTATCAATTTTATCAACAATGATGATGATGATGATGATGATGATG 960  
QY 961 AACCAATTTGAATTTGTCATTCGAGTTATTAATCAAAATGTTAAATTAATTTACGTA 1020  
Db 961 AACCAATTTGAATTTGTCATTCGAGTTATTAATCAAAATGTTAAATTAATTTACGTA 1020  
QY 1021 AAGAAGTGA 1029  
Db 1021 AAGAAGTGA 1029



RESULT 14  
US-08-827-356-1376  
; Sequence 1376, Application US/08827356  
; GENERAL INFORMATION:  
; APPLICANT: George H. Shimer, Jr.  
; APPLICANT: George H. Miller  
; APPLICANT: Robert S. Hare  
; APPLICANT: Karen J. Shaw  
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS RELATED  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS  
; NUMBER OF SEQUENCES: 5574  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Schering-Plough Corporation  
; STREET: 2000 Galloping Hill Road  
; CITY: Kenilworth  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07033-0530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentlin  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/827,356  
; FILING DATE: 01-APR-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/014,477  
; FILING DATE: 01-APR-1996  
; APPLICATION NUMBER: 60/016,743  
; FILING DATE: 02-MAY-1996  
; APPLICATION NUMBER: 60/020,016  
; FILING DATE: 14-JUN-1996  
; INFORMATION FOR SEQ ID NO: 1376:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1119 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Staphylococcus aureus  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: 1...1119  
; US-08-827-356-1376

Query Match 97.8%; Score 1006.4; DB 12; Length 1119;  
Best Local Similarity 99.1%; Pred. No. 9.4e-219;  
Matches 1023; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

QY 1 ATGCATATGCGATTCACCTTGCATTAATGTACAAAGTCACAAAGCGTGTAAATCCACCC 60  
DB 87 ATGCATATGCGATTCACCTTGCATTAATGTACAAAGTCACAAAGCGTGTAAATCCACCC 146  
QY 61 GTTGGGGCTGTGTAGTATGAGTAGATTTGTGTATGGTGCACACTTGAGAAAA 120  
DB 147 GTTGGGGCTGTGTAGTATGAGTAGATTTGTGTATGGTGCACACTTGAGAAAA 206  
QY 121 GGTGCAAGCATGCGAGAGTTCAGCACTTGATATGCGACACAAAGTCTGAAGTGGC 180  
DB 207 GGTGCAAGCATGCGAGAGTTCAGCACTTGATATGCGACACAAAGTCTGAAGTGGC 266  
QY 181 ACGATTATATATAGCTTAGAGCCATGTAGTCAATTTTGGTCAACACACCCGTGTAAAC 240  
DB 267 ACGATTATATATAGCTTAGAGCCATGTAGTCAATTTTGGTCAACACACCCGTGTAAAC 326  
QY 241 AAAATTATGTATGATAGTACGAAAGTAGATATACGCAACAAAGACAAATTCGTTAGAC 300

DB 327 AAAATTATGTATGATAGTACGAAAGTAGATATACGCAACAAAGACAAATTCGTTAGAC 386  
QY 301 ACACATGCTGATGAGACGTTACGGGCTCACGGATTTAGAGTTGAATCGTTGATGAGAA 360  
DB 387 ACACATGCTGATGAGACGTTACGGGCTCACGGATTTAGAGTTGAATCGTTGATGAGAA 446  
QY 361 CGGGCATACACATTTATACCAAGACTTTTAAAGCAAAAGCAAGCAACTGCCCAAAAT 420  
DB 447 CGGGCATACACATTTATACCAAGACTTTTAAAGCAAAAGCAAGCAACTGCCCAAAAT 506  
QY 421 ACAGTGAAGTATCTCGAAGTTAGATGTAACCAAGCAAGTATATGACCAACTGCA 480  
DB 507 ACAGTGAAGTATCTCGAAGTTAGATGTAACCAAGCAAGTATATGACCAACTGCA 566  
QY 481 TGCAATTAACAAAGAGTTAAACAGATGCTATTAAGTTAAGCATGACACGACGCA 540  
DB 567 TGCAATTAACAAAGAGTTAAACAGATGCTATTAAGTTAAGCATGACACGACGCA 626  
QY 541 GTGTTAAGTGAAGACGTACAGTTGAATTAAGATGATCCACAAATATACAGTATTCAA 600  
DB 627 GTGTTAAGTGAAGACGTACAGTTGAATTAAGATGATCCACAAATATACAGTATTCAA 686  
QY 601 GATGAAAAAACCCCTTAAAGATATTTGTCTAAGTCTGGGAATTTGATTTATTCAG 660  
DB 687 GATGAAAAAACCCCTTAAAGATATTTGTCTAAGTCTGGGAATTTGATTTATTCAG 746  
QY 661 CAATTTATCAAGATGATATCAACACCAATTTGATATATCAATCAATTTAAACA 720  
DB 747 CAATTTATCAAGATGATATCAACACCAATTTGATATATCAATCAATTTAAACA 806  
QY 721 AGCAATCAACACATATGAAATTTATTAAGTGAAGTCTGATTTAAACAATTCCT 780  
DB 807 AGCAATCAACACATATGAAATTTATTAAGTGAAGTCTGATTTAAACAATTCCT 866  
QY 781 CACAATTTATTAAGAGAGAGTGAAGTCTGATGAGGACGACGACCAACCACTACT 840  
DB 867 CACAATTTATTAAGAGAGAGTGAAGTCTGATGAGGACGACGACCAACCACTACT 926  
QY 841 TCAGAAATTCGCAATCT--ATTATATAGATGATTTATTCATATATGCCCCGAAATTA 897  
DB 927 TCAGAAATTCGCAATCTATTAATATAGATGATTTATTCATATATGCCCCGAAATTA 986  
QY 898 ATGGCGGATCTGGAATTTATCAATTTTATCAACAAGATGATGATGATCCAGAT 957  
DB 987 ATGGCGGATCTGGAATTTATCAATTTTATCAACAAGATGATGATGATCCAGAT 1046  
QY 958 GCGAACCAATTTGAATTTGATTCATTCGAGTTATTAATCAAAATGTAATTAATCTTA 1017  
DB 1047 GCGAACCAATTTGAATTTGATTCATTCGAGTTATTAATCAAAATGTAATTAATCTTA 1106  
QY 1018 CGAAGAGAGTGA 1029  
DB 1107 CGAAGAGAGTGA 1118

RESULT 15  
US-09-611-529-2079  
; Sequence 2079, Application US/09611529  
; GENERAL INFORMATION:  
; APPLICANT: George H. Shimer, Jr.  
; APPLICANT: George H. Miller  
; APPLICANT: Robert S. Hare  
; APPLICANT: Karen J. Shaw  
; TITLE OF INVENTION: Staphylococcus aureus Related Compositions and Methods  
; FILE REFERENCE: 1034/1C963051  
; CURRENT APPLICATION NUMBER: US/09/611,529  
; CURRENT FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/417,811  
; PRIOR FILING DATE: 1999-10-14  
; PRIOR APPLICATION NUMBER: US 09/353,718  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: US 09/266,557

1 PRIOR FILING DATE: 1999-03-11  
2 PRIOR APPLICATION NUMBER: US 09/266,556  
3 PRIOR FILING DATE: 1999-03-11  
4 PRIOR APPLICATION NUMBER: US 09/266,555  
5 PRIOR FILING DATE: 1999-03-11  
6 PRIOR APPLICATION NUMBER: US 09/266,542  
7 PRIOR FILING DATE: 1999-03-11  
8 PRIOR APPLICATION NUMBER: US 09/266,541  
9 PRIOR FILING DATE: 1999-03-11  
10 PRIOR APPLICATION NUMBER: US 09/037,934  
11 PRIOR FILING DATE: 1998-03-10  
12 PRIOR APPLICATION NUMBER: US 09/036,720  
13 PRIOR FILING DATE: 1998-03-06  
14 PRIOR APPLICATION NUMBER: US 09/036,338  
15 PRIOR FILING DATE: 1998-03-06  
16 PRIOR APPLICATION NUMBER: US 09/036,334  
17 PRIOR FILING DATE: 1998-03-06  
18 PRIOR APPLICATION NUMBER: US 09/036,221  
19 PRIOR FILING DATE: 1998-03-06  
20 PRIOR APPLICATION NUMBER: US 09/036,137  
21 PRIOR FILING DATE: 1998-03-06  
22 PRIOR APPLICATION NUMBER: US 09/036,082  
23 PRIOR FILING DATE: 1998-03-06  
24 PRIOR APPLICATION NUMBER: US 09/036,081  
25 PRIOR FILING DATE: 1998-03-06  
26 PRIOR APPLICATION NUMBER: US 09/036,079  
27 PRIOR FILING DATE: 1998-03-06  
28 PRIOR APPLICATION NUMBER: US 09/035,913  
29 PRIOR FILING DATE: 1998-03-06  
30 PRIOR APPLICATION NUMBER: US 09/035,744  
31 PRIOR FILING DATE: 1998-03-06  
32 PRIOR APPLICATION NUMBER: US 08/827,356  
33 PRIOR FILING DATE: 1997-04-01  
34 PRIOR APPLICATION NUMBER: US 08/831,156  
35 PRIOR FILING DATE: 1997-04-01  
36 PRIOR APPLICATION NUMBER: US 60/014,477  
37 PRIOR FILING DATE: 1996-04-01  
38 PRIOR APPLICATION NUMBER: US 60/016,743  
39 PRIOR FILING DATE: 1996-05-02  
40 PRIOR APPLICATION NUMBER: US 60/020,016  
41 PRIOR FILING DATE: 1996-06-14  
42 NUMBER OF SEQ ID NOS: 7451  
43 SEQ ID NO 2079  
44 LENGTH: 1119  
45 TYPE: DNA  
46 ORGANISM: *Staphylococcus aureus*  
47 US-09-611-529-2079

Query Match 97.8%; Score 1006.4; DB 23; Length 1119;

Best Local Similarity 99.1%; Pred. No. 9,4e-219;

Matches 1023; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

QY 1 ATGATATATCGATTCACATTCGCAATATGCTACAGGTCAAGAGTGTATTCACCC 60  
DB 88 atgataatcgatctcaactctgcaaatatgctacaggtcaagagtgtaataccccc 147  
QY 61 GTTGCGCCCTGTTAGTAAAGAGTGTGATTTGTTGATTTGTCACACTTGAGAAA 120  
DB 148 gttagcgctgtgtagtaataagagtgatgtgtgtgtgtgtgtgtgtgtgtgtgt 207  
QY 121 GGTGCAAGAGTGGGAGGTTCAAGCACTTGATATGTCACAAATAATGCTGAAGTGG 180  
DB 208 ggtgcaagagtggtgaggtgtcaagcttgatatagtgcaacaataatgtcgaagtg 267  
QY 181 ACGATTATATAGTACAGCCATGTAGTATTTGTTTACACACACCTGTGTAAAC 240  
DB 268 acgattatatactagcttagagccatgtagtaatttgttcaacacacacctgtgtaac 327  
QY 241 AAAATTATATGATTAAGATAGCAAAAGTACTATACCAACAAAGCAATTCGTTGAC 300  
DB 328 aaaattatctatgttaagataagcaaaagtgatatacgaacaagaacatctgttagac 387

QY 301 ACACATGCTGATGAGACGTTACGGGCTCACGATATTGAGTTGATGCTGATGATGAA 360  
DB 388 acacatgtgtatgtagacgtttacgggtccacgcttagtgatgtgtgtgtgtgtgtgt 447  
QY 361 CGGGCATTCACAAATTTATACCAAGACTTTTAAAGCAAAAGCAAAAGCAATTCGCA 420  
DB 448 cgggcattcacatatacccaagactttttaaagcaaaagcaaacatctgccaaatlt 507  
QY 421 ACAGTGAAGATATCTGCAAGTTTGAATGATTAACAGCAATGATTAATGCAAGTCAA 480  
DB 508 acagtgaaagtatctgcgaagtttagatgttaaacagcaagatgataagcaaaagtc 567  
QY 481 TGGATTACTAACAAGAGTTAAACAGATGCTATTAAGTTAAACATTCACAGACGCA 540  
DB 568 tggattactaaccaagaggtttaaaccaagatgtctataagtaagacatcgacagca 627  
QY 541 GTGTTACTGGAACAGTACGTTGATTAATGATATGCAATTTACTACAGTATTCAA 600  
DB 628 gtgttaacttgaagagcagctacagtttagatgtatccacaatactacacagatlt 687  
QY 601 GATGAAAAAACCTATTAAGATATATGCTAAGTCTGGAGATATTCATTTAATCAG 660  
DB 688 gatgaaaaaacctataaagaagtaatatgtctgaagctcggaatatcttaacttaac 747  
QY 661 CAAATTTATCAAGATGATCAACACCAATTTGATATATCTGAAAAATCCAAATTTACA 720  
DB 748 caaattatcaagatgtaatcaacacaaatttgatatactgtgaataatccaaatttaca 807  
QY 721 AGCAATTAACACATATTTGAATTTATTTACTTGAAGTCTGTGATTTAACAACATTC 780  
DB 808 agcaattacaacacacatctgaataatcttaacttgaagctgtgtgtgtgtgtgtgt 867  
QY 781 CACAATTTATTAAGAGAGAGTGTGAACCTTGCTAGTCGAGCGAGCGCAACCACTACT 840  
DB 868 cacaattataataaagagaggtgtgaacttgcgtcgagcgaggtccaaacactact 927  
QY 841 TCAGAAATCTCCATCT--ATTATATAGATGAATTTATCTATTTATGCCCCGAAATTA 897  
DB 928 tcgaatctccaatctcaattatataagatgaattatctctatltgtcccgaaatta 987  
QY 898 ATTGCGGATGTGAAATTTATCAATTTTATCAAAACAATGATGTGATTTAGATACCA 957  
DB 988 attgcgagatctggaattatcaatttatacaacaatgtatgtatgagatccagat 1047  
QY 958 GCGAACCAATTTGAATTTGATTCGAGTTATTAATCAAAATGTTAAATTAACCTTTA 1017  
DB 1048 gcgaaaccaatttgaattgttcaattccgagttatcaataacaaatgtuaatttaactt 1107  
QY 1018 CGAAAGAAGTGA 1029  
DB 1108 cgaagaagtgta 1119

Search completed: May 7, 2001, 12:41:33  
Job time: 7204 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 7, 2001, 11:48:14 : Search time 79.87 Seconds  
(without alignments)  
4404.654 Million cell updates/sec

Title: US-09-376-633-1

Perfect score: 1029

Sequence: 1 ATGATTTATGCGATTCACACT.....TAACCTTAGCAAGAAGTCA 1029

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 246011 seqs, 170942543 residues

492022

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_NA\_New.\*

1: /cgn2\_6/ptodata/1/pna/PCT\_NEW\_COMB.seq.\*

2: /cgn2\_6/ptodata/1/pna/US06\_NEW\_COMB.seq.\*

3: /cgn2\_6/ptodata/1/pna/US07\_NEW\_COMB.seq.\*

4: /cgn2\_6/ptodata/1/pna/US08\_NEW\_COMB.seq.\*

5: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq.\*

6: /cgn2\_6/ptodata/1/pna/US60\_NEW\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1000	97.2	7588	4 US-08-956-171C-142	Sequence 142, App
2	52.4	5.1	1250	5 US-09-739-449-1859	Sequence 1859, App
3	50.8	4.9	757	5 US-09-739-449-2286	Sequence 2286, App
4	50	4.9	70732	6 US-60-248-505-288	Sequence 288, App
5	49.8	4.8	1215	5 US-09-739-449-6666	Sequence 6666, App
6	49.8	4.8	506113	5 US-09-739-449-218	Sequence 218, App
7	47.8	4.6	518	5 US-09-833-381-1393	Sequence 1393, App
8	47.2	4.6	723	5 US-09-739-449-8012	Sequence 8012, App
9	47.2	4.6	70732	6 US-60-248-505-288	Sequence 288, App
10	47.2	4.6	70732	6 US-60-248-505-288	Sequence 288, App
11	47.2	4.6	70732	6 US-60-248-505-288	Sequence 288, App
12	47.2	4.6	70732	6 US-60-248-505-288	Sequence 288, App
13	46.2	4.5	776	5 US-09-739-449-1257	Sequence 1257, App
14	46.2	4.5	776	5 US-09-739-449-1257	Sequence 1257, App
15	44.2	4.3	902	5 US-09-739-449-1929	Sequence 1929, App
16	43.8	4.3	798	5 US-09-739-449-2641	Sequence 2641, App
17	43.8	4.3	798	5 US-09-739-449-2641	Sequence 2641, App
18	43.2	4.2	543	5 US-09-739-449-8029	Sequence 8029, App
19	43.2	4.2	543	5 US-09-739-449-8029	Sequence 8029, App
20	43.2	4.2	1131	5 US-09-739-449-1953	Sequence 1953, App
21	42.6	4.1	780	5 US-09-739-449-451	Sequence 451, App
22	42.6	4.1	1006	5 US-09-739-449-621	Sequence 621, App
23	42.2	4.1	1040	5 US-09-739-449-2287	Sequence 2287, App
24	42.2	4.1	834	5 US-09-739-449-2256	Sequence 2256, App
25	41.4	4.0	43546	6 US-60-254-168-38	Sequence 38, App
26	41	4.0	1418	5 US-09-739-449-1613	Sequence 1613, App
27	40.8	4.0	756	5 US-09-739-449-870	Sequence 870, App

C 28	40.8	4.0	787	5 US-09-739-449-2501	Sequence 2501, App
C 29	40.8	4.0	1343	5 US-09-801-833-6840	Sequence 6840, App
C 30	40.6	3.9	780	5 US-09-739-449-451	Sequence 451, App
C 31	40.6	3.9	910	5 US-09-739-449-2366	Sequence 2366, App
C 32	40.6	3.9	9834	4 US-08-956-171C-37	Sequence 37, App
C 33	40.4	3.9	830	5 US-08-739-449-370	Sequence 370, App
C 34	40.2	3.9	3931	4 US-08-956-171C-342	Sequence 342, App
C 35	40.2	3.8	69513	6 US-60-248-823-16	Sequence 16, App
C 36	39.6	3.8	461	5 US-09-737-223-15443	Sequence 15443, App
C 37	39.6	3.8	325507	6 US-60-248-505-267	Sequence 267, App
C 38	39.4	3.8	786	5 US-09-739-449-1368	Sequence 1368, App
C 39	39.2	3.8	757	5 US-09-739-449-2500	Sequence 2500, App
C 40	39	3.8	711	5 US-09-739-449-77	Sequence 77, App
C 41	39	3.8	909	5 US-09-739-449-636	Sequence 636, App
C 42	39	3.8	5136	4 US-08-956-171C-332	Sequence 332, App
C 43	38.8	3.8	778	5 US-09-739-449-2435	Sequence 2435, App
C 44	38.8	3.8	98573	6 US-60-248-505-542	Sequence 542, App
C 45	38.6	3.8	4358	5 US-09-308-453-1	Sequence 1, App

#### ALIGNMENTS

RESULT 1  
US-08-956-171C-142/c  
Sequence 142, Application US/08956171C  
GENERAL INFORMATION:  
APPLICANT: Charles Kunsch  
Gill H. Choi  
Patrick S. Dillon  
Craig A. Rosen  
Steven C. Barash  
Michael R. Fannon  
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
NUMBER OF SEQUENCES: 5255  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/956,171C  
FILING DATE: 20-Oct-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/009,861  
FILING DATE: January 5, 1996  
APPLICATION NUMBER: 08/781,986  
FILING DATE: January 3, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Hoover, Kenley K.  
REGISTRATION NUMBER: 40,302  
REFERENCE/DOCKET NUMBER: PB248P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 610-5790  
TELEFAX: (301) 309-8439  
INFORMATION FOR SEQ. ID NO: 142:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7588 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 142:  
US-08-956-171C-142















**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 7, 2001, 10:35:04 ; Search time 1372.91 Seconds  
(Without alignments)  
11053.563 Million cell updates/sec

Title: US-09-376-633-1  
Sequence: 1 ATGGATTATCGATTCAACT.....TAACCTTACGAAGAAGTGA 1029

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1283235 seqs, 7373929652 residues

Total number of hits satisfying chosen parameters: 2566470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl: \*  
1: gb\_ba1: \*  
2: gb\_ba2: \*  
3: gb\_ba3: \*  
4: gb\_in1: \*  
5: gb\_in2: \*  
6: gb\_in3: \*  
7: gb\_om: \*  
8: gb\_ov: \*  
9: gb\_pat1: \*  
10: gb\_pat2: \*  
11: gb\_ph: \*  
12: gb\_p11: \*  
13: gb\_p12: \*  
14: gb\_p13: \*  
15: gb\_p14: \*  
16: gb\_ba1: \*  
17: gb\_ba2: \*  
18: gb\_fun: \*  
19: gb\_htgo\_hum: \*  
20: gb\_htgo\_inv: \*  
21: gb\_htgo\_rod: \*  
22: gb\_htg\_hum1: \*  
23: gb\_htg\_hum2: \*  
24: gb\_htg\_hum3: \*  
25: gb\_htg\_hum4: \*  
26: gb\_htg\_hum5: \*  
27: gb\_htg\_hum6: \*  
28: gb\_htg\_hum7: \*  
29: gb\_htg\_hum8: \*  
30: gb\_htg\_inv1: \*  
31: gb\_htg\_inv2: \*  
32: gb\_htg\_rod: \*  
33: gb\_htg\_rod: \*  
34: gb\_hum1: \*  
35: gb\_hum2: \*  
36: gb\_hum3: \*  
37: gb\_hum4: \*  
38: gb\_hum5: \*  
39: gb\_hum6: \*  
40: gb\_hum7: \*  
41: gb\_in: \*  
42: gb\_om: \*  
43: gb\_ot: \*

44: em\_ov: \*  
45: em\_pat: \*  
46: em\_ph: \*  
47: em\_pl: \*  
48: em\_ro: \*  
49: em\_sts: \*  
50: em\_sy: \*  
51: em\_un: \*  
52: em\_v1: \*  
53: gb\_sts1: \*  
54: gb\_sts2: \*  
55: gb\_sts3: \*  
56: gb\_sy: \*  
57: gb\_un: \*  
58: gb\_v11: \*  
59: gb\_v12: \*  
60: gb\_hvg1: \*  
61: gb\_hvg2: \*  
62: gb\_hvg3: \*  
63: gb\_hvg4: \*  
64: gb\_hvg5: \*  
65: gb\_hvg6: \*  
66: gb\_hvg7: \*  
67: gb\_hvg8: \*  
68: gb\_hvg9: \*  
69: gb\_hvg10: \*  
70: gb\_hvg11: \*  
71: gb\_hvg12: \*  
72: gb\_hvg13: \*  
73: gb\_hvg14: \*  
74: gb\_hvg15: \*  
75: gb\_hvg16: \*  
76: gb\_hvg17: \*  
77: gb\_hvg18: \*  
78: gb\_hvg19: \*  
79: gb\_hvg20: \*  
80: gb\_hvg21: \*  
81: gb\_hvg22: \*  
82: gb\_hvg23: \*  
83: gb\_hvg24: \*  
84: gb\_hvg25: \*  
85: gb\_pr1: \*  
86: gb\_pr2: \*  
87: gb\_pr3: \*  
88: gb\_pr4: \*  
89: gb\_pr5: \*  
90: gb\_pr6: \*  
91: gb\_pr7: \*  
92: gb\_pr8: \*  
93: gb\_pr9: \*  
94: gb\_rol: \*  
95: gb\_rol2: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1029	100.0	3336	9 AR069579	AR069579 Sequence
2	516.8	50.2	3227	2 AF269646	AF269646 Staphyloc
3	516.8	50.2	3618	2 AF269345	AF269345 Staphyloc
4	340.6	33.1	2902	2 AF270130	AF270130 Staphyloc
5	155	15.1	3721	2 AF269926	AF269926 Staphyloc
6	148	14.4	300550	2 AP001512	AP001512 Bacillus
7	121.2	11.8	293181	2 AP001119	AP001119 Buchnera
8	120.4	11.7	6006	2 BSRIB	X51510 B. subtilis
9	120.4	11.7	28206	2 BACDIA	L09228 Bacillus su
10	120.4	11.7	218470	2 BSDB0013	Z59116 Bacillus su
11	116.4	11.3	14364	1 AE000675	AE000675 Aquifex a

```
12 112.2 10.9 4286 2 BARIBGENS X95955 B.amyloliqu
13 111 10.8 175936 84 SPNE01908 AL449930 Streptococ
14 105.8 10.3 11167 1 AE006333 Lactococc AE006333
15 95.2 9.3 18073 3 U32775 Haemophilus U32775
16 91.4 8.9 11695 1 AE004298 Vibrio ch AE004298
17 85.8 8.3 10166 1 AE006112 Pasteurel AE006112
18 83.4 8.1 12811 1 AE001819 Thermotog AE001819
19 79 7.7 5451 3 TNE9832 Thermotog AJ009832
20 78.6 7.6 2949 2 ECNUSB X64395 E.coli gene
21 78.6 7.6 11692 1 AE000148 AE000148 Escherich
22 78.6 7.6 139818 2 ECOR2664 U82664 Escherichia
23 77.4 7.5 105795 3 SYCCPBC U27202 Actinobacil
24 76.2 7.4 4312 2 AP027202 U27202
25 75.4 7.3 2759 1 AF002857 Shigella AF002857
26 75.4 7.3 10323 1 AE005220 Escherich AX067466
27 71.2 6.9 269223 10 AX067466 Sequence
28 67.8 6.6 11456 1 AE002377 Chlamydia AE002277
29 67.6 6.6 18759 1 AE002257 Chlamydia AE002257
30 67.6 6.6 20386 1 AE001667 Chlamydia AE001667
31 67.6 6.6 325865 2 AP002548 Chlamydia AP002548
32 66 6.4 11544 1 AE001343 Chlamydia AE001343
33 65.4 6.4 14376 1 AE002531 Neisseria AE002531
34 65.4 6.4 349980 9 AX044034 AX044034 Sequence
35 62.4 6.1 349061 3 NMA223491 AL162753
36 60.4 5.9 2307 2 BHE132928 Neisseria AJ132928
37 60.4 5.9 85779 14 SCE011856 Bartonell AJ011856
38 60.2 5.9 86827 6 PEMALP5 AL034556
39 59.8 5.8 964 54 CNS07E8R AL034556
40 59.6 5.8 11404 1 AE003934 Al441457
41 58 5.6 60604 68 AC023466 Xylella f AC023466
42 56.4 5.5 150236 91 HS101G11 AC023466
43 56.4 5.5 205429 60 AC005506 Human DNA AC005506
44 55.6 5.4 7218 10 166494 Sequence
45 55.6 5.4 162160 63 AC013817 Homo sapi AC013817
```

## ALIGNMENTS

```
RESULT 1
AR069579 3336 bp DNA PAT 18-FEB-2000
LOCUS AR069579
DEFINITION Sequence 7 from patent US 5891672.
ACCESSION AR069579
VERSION AR069579.1 GI:7220467
KEYWORDS
SOURCE
ORGANISM
UNKNOWN.
```

## REFERENCE

1 (bases 1 to 3336)  
Wang, M., Ward, J. M., Warren, R. L., Lloyd, Nicholas, R. Oakley, Palmer, L., Marle, Pratt, J. M., Knowles, D., Justincharles, Loretto, M., Arthur, Mooney, J., Black, M., Terence, Burnham, M., Karl, Russell, Debouck, C., Fedon, J. Craig, Hodgson, J. Edward, Javoriski, D. Dee, Reichard, R. Winfield, Rosenberg, M., Traini, C., Michael and Zhong, Y. X.  
Polynucleotides encoding GTP cyclohydrolase II (riba)  
Patent: US 5891672-A 7 06-APR-1999;  
Location/Qualifiers  
1..3336  
source

BASE COUNT 1202 a 459 c 678 g 997 t  
ORIGIN

Query Match 100.0% Score 1029; DB 9; Length 3336;  
Best Local Similarity 100.0%; Pred. No. 6.3e-179;  
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGGATTATGCGATTCACGCTTGAATATGATGATACAGGTCACAGGTTTATCCACCC 60
DB 1 ATGGATTATGCGATTCACGCTTGAATATGATGATACAGGTCACAGGTTTATCCACCC 60
```

```
QY 61 GTTGGCGCTGTGTGACTATATAGAGTAGAGTGTGTGTTGTGTGCGACCTTGAGAAA 120
DB 61 GTTGGCGCGCTGTGTGACTATATAGAGTAGAGTGTGTGTTGTGTGCGACCTTGAGAAA 120
QY 121 GGTGACCAACGATCGGAGGTTCAAGCAGCTGTGATGAGGCAACCAAAATGCTAAGTGG 180
DB 121 GGTGACCAACGATCGGAGGTTCAAGCAGCTGTGATGAGGCAACCAAAATGCTAAGTGG 180
QY 181 ACAGATTATATACGTTAAGGCAAGTAGTCACTTTTGGTTTCAACACCACTGTGTAAAC 240
DB 181 ACAGATTATATACGTTAAGGCAAGTAGTCACTTTTGGTTTCAACACCACTGTGTAAAC 240
QY 241 AAAATTATGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
DB 241 AAAATTATGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 301 ACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
DB 301 ACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
QY 361 CGGCGATCACAATTTATACCAAGCTTTTAAAGCAAAAGCAAAAGCAAAAGCAAAAGT 420
DB 361 CGGCGATCACAATTTATACCAAGCTTTTAAAGCAAAAGCAAAAGCAAAAGCAAAAGT 420
QY 421 ACAGTGAAGATATCTGCAAGTTAGATGATGATGATGATGATGATGATGATGATGATGAT 480
DB 421 ACAGTGAAGATATCTGCAAGTTAGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 481 TGGATTACATCAACCAAGGTTTAAACAGATGCTATTAAGTTAAGCATGATGATGATGAT 540
DB 481 TGGATTACATCAACCAAGGTTTAAACAGATGCTATTAAGTTAAGCATGATGATGATGAT 540
QY 541 GGTGTTAAGTGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 600
DB 541 GGTGTTAAGTGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 600
QY 601 GATGAAAAAACCCTATATAAGTATATGCTAAGTCTGGAATTTATTTTAATCAG 660
DB 601 GATGAAAAAACCCTATATAAGTATATGCTAAGTCTGGAATTTATTTTAATCAG 660
QY 661 CAAATTTATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
DB 661 CAAATTTATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
QY 721 ACACATCAACACATATTTGAATTTATTTAAGTCTGGAATTTATTTTAATCAG 780
DB 721 ACACATCAACACATATTTGAATTTATTTAAGTCTGGAATTTATTTTAATCAG 780
QY 781 CACATTTTATATAAAGAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 781 CACATTTTATATAAAGAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 841 TCGAATTTCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
DB 841 TCGAATTTCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
QY 901 GCGGATCTGGAATTTATATCAATTTATATCAACCAATGATGATGATGATGATGATGATGAT 960
DB 901 GCGGATCTGGAATTTATATCAATTTATATCAACCAATGATGATGATGATGATGATGATGAT 960
QY 961 AACCAATTTGAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
DB 961 AACCAATTTGAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
QY 1021 AAGAGTGA 1029
DB 1021 AAGAGTGA 1029
```

RESULT 2  
AF269646/c 3227 bp DNA BCT 01-AUG-2000  
LOCUS AF269646/c  
DEFINITION Staphylococcus epidermidis strain SRI clone step.1015d10 genomic





```

source
1. .3618
/organism="Staphylococcus epidermidis"
/strain="SRI"
/db_xref="taxon:1282"
/clone="step_1002d07"

BASE COUNT      1099 a      709 c      518 g      1292 t
ORIGIN

```

Query Match	50.2%;	Score 516.8;	DB 2;	Length 3618;
Best Local Similarity	69.5%;	Pred. No. 2.8e-85;		
Matches 717; Conservative	0;	Mismatches 312;	Indels 3;	Gaps 1.

OY	1	ATGATATATGCGATTCACAACTTGGAAATATATGCTTCAGACGTGCAACAAAGGTTTAAATCCACC	60
Db	2701	ATGATATATGCGATTCACAACTTGGAAATATATGCTTCAGACGTGCAACAAAGGTTTAAATCCACC	2612
OY	61	GTTGCGCGCTGTTGACTTAATYGAAGGTAGAGATTGTTGCTATTGGTCACACCTTGAGAAA	120
Db	2641	GTAGAGATCCGCTGTTGTTAAAAACCGTAGAGATTGTGTAGGTTTGTGTGACATTTAAAG	2582
OY	121	GGTACAAACGATGGGGAGGTTTCAAGCACTTGATATGTGGCAACAAATGTGTAAAGTGTG	180
Db	2581	GGAAATTAACATGCGCGAAGTACAGCTATTGGAATGTGACGTTTAAATCCCAAGCTCT	2522
OY	181	ACGATTTATATATACGTTAGAGCCCATGTAGTCATTTTGGTTCACACCAACCCGTGTATAC	240
Db	2521	ACCATATACGTTTCATTATGAAACCTTGACACACCATGGTTCAACACCACTGTGTGAT	2462
OY	241	AAATTTATTCATTTCTAAGATACGAAATAGTATACGCAACAAAGACAAATTCGTTAGAC	300
Db	2461	AAATATCATTAACCGGCAATATCTAAAGTCATCTATGTGCTTAAAGATATCACTTTAGTA	2402
OY	301	ACACATGCGTATGAGACGTTTACGGGCTCACGCGTATGAGATTTGAATGGTGTGATGATGA	360
Db	2401	AGTAAGGTCGACGATTTTGAGAGAAGCGGTATAGAGTTGAATTTCAATATATATCA	2342
OY	361	CGGCGATTCACAAATTTATPCAAAGCTTTTTTAAAGCAAAAGCAACCACTGCCACAATTT	420
Db	2341	AAATGACACTGCATTTATTAACCGTGACTTTTTTACTGCTTAAAGAAAGCAAGGTTCCAAAGTA	2282
OY	421	ACAGTGAAGATATCTGCAAAATTTAGATGTGTAAACCAAGCAATGATATGACCAAGTCAA	480
Db	2281	ACGTAAAGGTCATCTATCTATCTAGATGTGTAAACCAAGCAACAGACTTTATATGAAGTAA	2222
OY	481	TGGATTTACTAACAAAGGTTTAAACAAGATGCTCTATAGTTAAGACATGACACAGCGCA	540
Db	2221	TGCTATTAACAAACAAAGAGTTTAAAGAGATGTTTATCAATTTAAGACATGACAGATATGCA	2162
OY	541	GTTTAACTGGAAGACGTACAGTTGGAATTAGATGATCCAAATATAGTACACGTATTCAA	600
Db	2161	GTTTATTTACTGGGCGTAAACCAATTTGAACGCAACATCCATTTGTATACCAACGAGGTTCT	2102
OY	601	GATGGAACAAAAACCTATTAAGAGTAATATTTGCTAACTGGCAATATTCATTTTAATCAG	660
Db	2101	GATGGAAGACATCGAGATTTCAGTTATTTCTTTCTTAAGAAAGTCACTGATTTTATTCAA	2042
OY	661	CAAAATTTATCAAGTATGATATACCAACAATTTGGAATATATCTGAAAAATTCAAATTTAACA	720
Db	2041	CAAAATTTTAAAGCTACTGATCGATGAGATATAGATTTATACAGAAAATGAAAATTTAA	1982
OY	721	AGCAATCAACACATATATGAAATTTATTTACTTGAAGCTCTGTGATTTTAAACAATTTCT	780
Db	1981	ACAATTAAGGTTTATTTAAATATATATATATATAGTATTTGATATCAACGACATATTA	1922
OY	781	CACAATTTATTAATAAAGAGGATTGCAACTTGTCTAGTCGAGGCAAGTCCAACTACT	840
Db	1921	CAGACTTTATATCAAGAGGCTATTTGGAAACGTCTAGTGTGAGGCGAGGCCCAAAATTTACA	1862
OY	841	TCGAAATTCCTCAATCT---ATTATATAGATGAAATTTATCTCTATTTATGCCCCGAAATTA	897
Db	1861	TCCTCAATTTCTCAATCAACATCTTAATATGACTCTATTTATATATAGCCCCGAAATTA	1802

[illegible]

RESULT	4
AF270130	AF270130
LOCUS	2902 bp DNA
DEFINITION	BCT 01-AUG-2000 Staphylococcus epidermidis strain SKI clone step.1051a10 genomic sequence.
ACCESSION	AF270130
VERSION	AF270130.1
KEYWORDS	GI:9624037
SOURCE	.
ORGANISM	Staphylococcus epidermidis. Staphylococcus epidermidis.

	Query Match	33.1%	Score 340.6;	DB 2;	Length 2902;
	Best Local Similarity	69.0%	Pred. No. 4,8e-53;		
	Matches 466;	Conservative 0;	Mismatches 209;	Indels 0;	Gaps 0;
QY	1	ATGGATTATGCGATTCAACTTGCACAAATATGGTACAAAGGTCAACACAGGTTAATACACCC	60		
Db	2220	ATGGATGATGCGATTCAACTACGACAAATAATGGTAATGACAAACAGCGTTAATACACCA	2279		
QY	61	GTTGGCGGTGTGTAGTAAATGAAGGTGAGATGTGTGTATGTGGGCACACTGGAAAA	120		
Db	2280	GTAGATCCGATGTGTGTTAAACACGGTGGATGTAGGTTTAGTGTCACATTAAAAAAG	2339		
QY	121	GGTGCAAGCATGCGGAGGTTCAAGCACTTGATATGAGCAACAAATGCTGAAGGTGCG	180		
Db	2340	GGAGATAAACAATGCCGAAGTACAAACCTATGAAATGGCAGGTTAAATACCAAGGTGCT	2399		
QY	181	ACGATTATATTTACGTTAGAGACCATGTAGTCATTTTGGTTCAACACCACTGTGTAAAC	240		
Db	2400	ACCATATATCGTTTCAATTATGAAACCTTGGACACACCACTGTGTCAACACCACTTGTGTGAT	2459		

QY	241	AAATTTTGAATGTAGTAGAGCAAAAGTAGTACCAACAAGACAATTCGTTAGAC	300
Db	2460	AAATTCATTGAAGGGGCATATCTAAAGTCACTATGCTGTTAAAGATCTACTTTAGTA	2519
QY	301	ACACATGSGTAGTAGACGTTACGGGCTCAACGGTATTGAGTTGAATGGCTTGATGATGAA	360
Db	2520	AGTAGAGGTGACGATTTGACAGAAAGCTGGTATAGAGTTGGAATTTCAATATATATGA	2579
QY	361	CGGAGCATCACAAATTATTCACAGACTTTTAAAGCAAAAGCAACACACTCGCCAAATT	420
Db	2580	AATGCACTGCATTATTAACCGTGACTTTTACTGCTAAAGAAAGCAAGGTTCCGAAGTA	2639
QY	421	ACAGTGAAGATATCTGCCAACTTTAGTGTAAACAAGCAAGATATATGACAAAGTAA	480
Db	2640	ACTGTAAAGGTCATCTACTACTAGATGTAACGAACAAGCAACAGACTTTTAATGAAGTAG	2699
QY	481	TGGAATTCTAACAAAGGTTAAACAAGATGCTCTATAGTTAAGCAATCGACAGCA	540
Db	2700	TGGATTAACAACAAGAAAGTTAAAGAAAGTATTATCAATTATTAAGCAATGACATGATGCA	2759
QY	541	GTTGTAACGTGAAGACTACTAGTTGAATTAGATGATCCAAATTAATTAATACACGATTTCAA	600
Db	2760	GTTATTACTGGGCTAGAACCATCTGAAGCAGACATCTTGTATACCAACGAGGTTCT	2819
QY	601	GATGGAAGAAAAACCTATTAAGATATATTTGCTAAGTCTGGGAATATTCATTTTAATCG	660
Db	2820	GATGGAAGACATCCGATTCAGATTCTTTCTTAAGAAAGTCACTCGATTTTAATCAA	2879
QY	661	CAATTTATCAAGT 675	
Db	2880	CAATATATTTAAAGT 2894	

RESULT	5	AF269926/c	AF269926	3721 bp	DNA	BC1	01-AUG-2000
LOCUS		Staphylococcus epidermidis strain SRI clone step.1032d02 genomic sequence.	AF269926				
ACCESSION		AF269926					
VERSION		AF269926.1	GI:9623827				
KEYWORDS							
SOURCE		Staphylococcus epidermidis.					
ORGANISM		Staphylococcus epidermidis.					
REFERENCE		Bacteria; Firmicutes; Bacillus/Clostridium group; Staphylococcus.					
AUTHORS		1 (bases 1 to 3721) Kimmerly W.U., Taylor J.David, Nelsen A.J., Godlewski M.M., Rudino M.A., Nelson F.J., Rivers P.R., Tortunella-Miller I., Listenebee S., Ashanti C., Altschuller G., Mamo L., Shepherd N.S., Fuchs R., Fleming T., Guan X., Du L., Cain D.H., Miller G.S. and Furdon P.J.					
TITLE		Transposon-mediated sequencing of the Staphylococcus epidermidis genome					
JOURNAL		Unpublished					
REFERENCE		2 (bases 1 to 3721)					
AUTHORS		Taylor J.David, Kimmerly W.J., Nelsen A.J., Godlewski M.M., Rudino M.A., Nelson F.J., Rivers P.R., Tortunella-Miller I., Listenebee S., Ashanti C., Altschuller G., Mamo L., Shepherd N.S., Fuchs R., Fleming T., Guan X., Du L., Cain D.H., Miller G.S. and Furdon P.J.					
TITLE		Direct Submission					
JOURNAL		Submitted (22-MAY-2000) Departments of Genomic Sciences and Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore Drive, Research Triangle Park, North Carolina 27709-3398, USA					
FEATURES		Location/Qualifiers					
SOURCE		1..3721					
		/organism="Staphylococcus epidermidis"					
		/strain="SRI"					
		/db_xref="taxon:1282"					
		/clone="step.1032d02"					
BASE COUNT		1204 a 625 c 495 g 1397 t					
ORIGIN							

Query Match	15.1%	Score 155:	DB 2:	Length 3721:
Best Local Similarity	76.1%:	Pred. No. 4e-19:		
Matches 191:	Conservative	0:	Mismatches 60:	Indels 0:
			Gaps	0:
QY	1	ATGAGTTATGCGATTCACACTTGCATAATATATGACACAGGTCAACAGCGTTTAAATCACC	60	
Db	259	ATGATGATGATGCGATTCACACTAGCAAAAATGTGAATATGCAACAAACAGGTGTTAAATCCACCA	200	
QY	61	GTTGGCGCTGTTTGATGTTAATGAAAGTAGAGATGTTGGTATTTGTCACACTTGAAGAAA	120	
Db	199	GTAGATCCGTTGTTGTTAAAAACGTAAGATGTAGGTTAGTGTCACATTTAAAAANG	140	
QY	121	GCTGACAGCATGCCGAGCTTCAGCACTGTATATGACACAACAAATGCTGAGGTGCG	180	
Db	139	GGAGTTAAACATGCGCGAAGTAGCAACCTTATTTGAAAGTACGAGTTTAAATACCAAGGTGCT	80	
QY	181	ACGATTTATTTAGCTTAGAGCCATGTATGCTATTTTGGTTCAAACACACCCCTGTGTAC	240	
Db	79	ACCATATATCGTTTCATTTAGAACCTTGACACACACATGTGTTCAACACACACTTGTGTGCAT	20	
QY	241	AAATTTATGA	251	
Db	19	AAATCATTTGA	9	

RESULT	6
LOCUS	AP001512
DEFINITION	AP001512 300550 bp DNA BCT 10-JAN-2001
ACCESSION	Bacillus halodurans genomic DNA, section 6/14.
VERSION	AP001512 BA000004
KEYWORDS	AP001512.1 GI:10174030
SOURCE	.
ORGANISM	Bacillus halodurans DNA. Bacillus halodurans Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillus/Staphylococcus group; Bacillus.
REFERENCE	1 (sites)
AUTHORS	Takami,H., Nakasone,K., Hiramata,C., Takaki,Y., Masui,N., Fujii,F., Nakamura,Y., and Inoue,A. An improved physical and genetic map of the genome of alkaliphillic Bacillus sp. C-125
JOURNAL	Extremophiles 3 (1), 21-28 (1999) MEDLINE 99184645
REFERENCE	2 (sites)
AUTHORS	Takami,H., Nakasone,K., Ogasawara,N., Hiramata,C., Nakamura,Y., Masui,N., Fujii,F., Takaki,Y., Inoue,A. and Horikoshi,K. Sequencing of three lambda clones from the genome of alkaliphillic Bacillus sp. strain C-125
JOURNAL	Extremophiles 3 (1), 29-34 (1999) MEDLINE 99184646
REFERENCE	3 (sites)
AUTHORS	Takami,H., Takaki,Y., Nakasone,K., Hiramata,C., Inoue,A. and Horikoshi,K. Sequence analysis of a 32-kb region including the major ribosomal protein gene clusters from alkaliphillic Bacillus sp. strain C-125
JOURNAL	Biosci. Biotechnol. Biochem. 63 (2), 452-455 (1999) MEDLINE 99209008
REFERENCE	4 (sites)
AUTHORS	Takami,H., Masui,N., Nakasone,K. and Horikoshi,K. Replication origin region of the chromosome of alkaliphillic Bacillus halodurans C-125
JOURNAL	Biosci. Biotechnol. Biochem. 63 (6), 1134-1137 (1999) MEDLINE 99356711
REFERENCE	5 (sites)
AUTHORS	Takami,H., Takaki,Y., Nakasone,K., Sakiyama,T., Maeno,G., Sasaki,R., Hiramata,C., Fujii,F. and Masui,N. Genetic analysis of the chromosome of alkaliphillic Bacillus halodurans C-125
JOURNAL	Extremophiles 3 (3), 227-233 (1999) MEDLINE 99411980
REFERENCE	6 (sites)

AUTHORS	Takami, H.
TITLE	Genome analysis of facultatively alkaliphilic Bacillus halodurans C-125
JOURNAL	(1n) Extremophiles in deep-sea environments (Ed.):
REFERENCE	: 249-284; Springer-Verlag (1999)
AUTHORS	7 (sites)
TITLE	Takami, H. and Horikoshi, K.
JOURNAL	Reidentification of facultatively alkaliphilic Bacillus sp. C-125 to Bacillus halodurans
REFERENCE	Biosci. Biotechnol. Biochem. 63, 943-945 (1999)
AUTHORS	8 (sites)
TITLE	Takami, H., Nakasone, K., Takaki, Y., Maeno, G., Sasaki, R., Masui, N., Fuji, F., Hirama, C., Nakamura, Y., Ogasawara, N., Kuhara, S. and Horikoshi, K.
JOURNAL	Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and genomic sequence comparison with Bacillus subtilis
REFERENCE	Nucleic Acids Res. 28, 4317-4331 (2000)
TITLE	9 (sites)
JOURNAL	Takami, H. and Horikoshi, K.
REFERENCE	Analysis of the genome of an alkaliphilic Bacillus strain from an industrial point of view
AUTHORS	Extremophiles 4 (2), 99-108 (2000)
JOURNAL	20263314
REFERENCE	10 (sites)
AUTHORS	Nakasone, K., Masui, N., Takaki, Y., Sasaki, R., Maeno, G., Sakiyama, T., Hirama, C., Fuji, F. and Takami, H.
TITLE	Characterization and comparative study of the rrn operons of alkaliphilic Bacillus halodurans C-125
JOURNAL	Extremophiles 4 (4), 209-214 (2000)
REFERENCE	11 (bases 1 to 300550)
JOURNAL	Takami, H. and Takaki, Y.
REFERENCE	Submitted (22-MAR-2000) Hideyo Takami, Japan Marine Science and Technology Center, Deep-sea Microorganisms Research Group; 2-15 Natsushima, Yokosuka, Kanagawa 237-0061, Japan
TITLE	(E-mail: takami@jamstec.go.jp; URL: <a href="http://www.jamstec.go.jp/jamstec/e/bio/DEEPSTAR/Research.html">http://www.jamstec.go.jp/jamstec/e/bio/DEEPSTAR/Research.html</a> ), Tel: 81-468-67-3895, Fax: 81-468-66-6364)
FEATURES	Location/Qualifiers
source	1. 300550
gene	/organism="Bacillus halodurans"
CDS	/db_xref="taxon:86665"
gene	/note="alkaliphile"
CDS	157..816
gene	/gene="BH1414"
CDS	157..816
gene	/gene="BH1414"
CDS	/note="BH1414 unknown"
gene	/codon_start=1
CDS	/transl_table=1
gene	/protein_id="BAB05133.1"
CDS	/db_xref="GI:10174031"
gene	/translation="MFEQGEQLLEILIKRQGLSFRKIAQALNTVGVKHYRKKH
gene	ELQDQKQAKPAKGEQQLGSGVPEVETILDPGGPTLYAFMPTSEOTRGVSYQDDE
gene	KMDQLKRLRLIVYTSILFPGSRANVYVDIHLDEPMNNMLIEHVEPCTYLIDGLVGI
gene	TPSGAFLPYVARSSTPIETPRQNDYINRKEVDYDMQIGLSEPKMLEHFSITYYEKT
gene	K"
gene	829..3600
gene	/gene="BH1415"
CDS	829..3600
gene	/gene="BH1415"
CDS	/note="BH1415 unknown conserved protein in others"
gene	/codon_start=1
CDS	/transl_table=1
gene	/protein_id="BAB05134.1"
CDS	/db_xref="GI:10174032"
gene	/translation="MREGQGSIVYHAHLPYVRHEDRLEERMLPEAMSEYIYLLNNA
gene	LEKPYKHAAYTISTPPYMEKLSPPVYOTRLNLHENTBQLKREKRTINQRTQNLV
gene	QFYKQREKKAIFLQMDRNLIGFSRLIMEEQCTLMTSAVTHAFPYLKKREKIRAG

[illegible]





[illegible]

## CDS

```

FLKEDEDIVCAIVEKKDSEVERIHRHSVNPNSHRHOGIGKOMMDALKHLFKTOVLVPN
ELTOSFEERCQGGDDDISYNN"
complement(5164..5688)
/note="ORF y (AA 1-174)"
/codon_start=1
/transl_table=11
/protein_id="CA35883.1"
/db_xref="GI:580916"
/db_xref="Swiss-Prot:P17617"
/translation="MPKAYIDYVEFHATPDYFECHIELEWEKDEPPKKRKYWG
FIOLAVLYHRRONTAGACRRLMANSIILQAEHRAVEDLGLDGRLELEWQSYEOI
ETSAVKSIIIPIDENKLEACRIECCRKKKTWQPSALNIFLIDKRLRLDRDVIIR
EREKEIERKRSRD"
BASE COUNT      1891 a      1167 c      1401 g      1547 t
ORIGIN

```

Query Match 11.7% Score 120.4 DB 2 Length 6006;  
 Best Local Similarity 51.0% Pred. No. 8.3e-13;  
 Matches 317; Conservative 0; Mismatches 296; Indels 9; Gaps 1;

```

OY 1 ATGATTTATGCGATTCACCTTCGAATATGTTACAGAGTCAACAGGTTTATCCACC 60
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1227 ATGAAGCGCGCCTTAGATCTTCGGAAGCAGCGGCAAGAGCAGACCCGAATCCATCCGCTC 1286
OY 61 GTTGGCGCTGTTGTAAGTAAGCAGTACGATTTGCTATTGCTGACACTTGAAGAAA 120
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1287 GTGGCGCTGTTGCTGAAGAGCAGCAAAATGTCGAAATGGCGCCCATTTAAATAT 1346
OY 121 GGTGCAAGCATGGGAGGTTTCAACACTTGTATGCGACACACAAATGCTGAAGTGG 180
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1347 GGTGAAGCTCATGAGAAAGTTATGTCATCCATATGCTGGAGCACATGCAAGGCTGCC 1406
OY 181 AGCATTTATATTACGTTAGAGCATGATGTCATTTGTTGTTCAACACCAACCTGTGTTAC 240
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1407 GACATTTACGTTACACTGCAACCGTACGCCATTTACGGAAAAAACCCGCAATGCGAGAA 1466
OY 241 AAAATTTATGATTTGAAGTACGCAAAAGTACATACCCACAAAGA-----CAAT 291
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1467 TTGATTTATCACTGATGTAAGAAAGTGTTCGTGGGATGAGATGCTTAATCCGCTT 1526
OY 292 TCGTTAGACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 351
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1527 GTGGCTGGAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1586
OY 352 GATGATGAACGGGATCACAATTTATACCAAGACTTTTAAAGCAAAAGCAAAAGCAACTG 411
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1587 CTGGCAGACCAAGCGGAGGCTGAATGAAATTTCTGCACTTTATGAGACAGGCGCTT 1646
OY 412 CCACAATTTACGTAAGTATCTGCAAGTTAGATGCTAATCAACAGCGAATGATATATGA 471
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1647 CCGTAGCTACGCTAAAGCGGCTGCCAGCCTTGACGGCAAGATGATGACGACAGGCTT 1706
OY 472 CAAGTCAATGATTTACTTAACAAGAGTTTAAACAAGATGCTTAAAGTTAGTACATGCA 531
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1707 GACACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1766
OY 532 CACGACGAGAGTGTAACTGGAAGACGTAAGTGAATAGATGATGATGATGATGATGATG 591
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1767 CACCAACGATTTTGTAGTGTGACAGTGAAGCCGACAAATCCGACGCTTAACTGCTC 1826
OY 592 CCGTATTCGAAGTGAAGAAACC 613
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1827 AGACTGCCGAATGTAACAAAC 1848
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```

```

RESULT 9
LOCUS      BACDIA      28206 bp      DNA
DEFINITION Bacillus subtilis spova to sera region.
ACCESSION  L09228
VERSION     L09228.1 GI:410114
KEYWORDS   3-dehydroquinase dehydratase; aroc gene; diaminiopimelate

```

## SOURCE

decarboxylase; lysa gene; penicillin-binding protein; peptidyl-prolyl isomerase; phosphoglycerate dehydrogenase; pglb gene; response regulator; response regulator kinase; riba gene; ribb gene; ribd gene; ribg gene; ribh gene; ribt gene; riboflavin biosynthesis operon; sera gene; signal peptidase; slps gene; spoa gene; spovaf gene.

## REFERENCE

1 (bases 1 to 28206)

Molecular cloning and analysis of nucleotide sequence of the Bacillus subtilis lysa gene region using B. subtilis phage vectors and a multi-copy plasmid, pUB110

Agric. Biol. Chem. 55 (6), 1615-1626 (1991)

91345841

## JOURNAL

2 (bases 1 to 28206)

Buchanan, C.E. and Ling, M.L.

Isolation and sequence analysis of dacB, which encodes a sporulation-specific penicillin-binding protein in Bacillus subtilis

J. Bacteriol. 174 (6), 1717-1725 (1992)

92193254

## MEDLINE

3 (bases 1 to 28206)

Sorokin, A., Zumbstein, E., Azavedo, V., Ehrlich, S.D. and Serror, P.

The organization of the Bacillus subtilis 168 chromosome region between the spova and sera genetic loci, based on sequence data

Mol. Microbiol. 10 (2), 385-395 (1993)

95020538

## FEATURES

1..28206

location/Qualifiers

/organism="Bacillus subtilis"

/strain="168"

/sub\_species="Marburg"

/db\_xref="taxon:1423"

1..1239

/gene="spovaf"

<1..1239

/gene="spovaf"

/codon\_start=1

/transl\_table=11

/protein\_id="AA67472.1"

/db\_xref="GI:410115"

/translation="VDIVENRLNAOYEKVTLDDETDOVLSGLAVIIVEGAFIT

DVRSYGRNPEEPDTEKVRGARDGFENIVNTALRRIRDERLRYKTKVERSK

TDLISCIYEDTADDELVEIYEKEIASIDVDLTMAADTVEEFIYNOGNPPLRYTE

RPDVAANVHSEGHVILIVDPSVYITPTLFLHVOHAEVROASVGTFLRMVRFPG

ILASTPEIWEFLFVLQDILPDMMKFTGKPKDHPHIILOITPTTPEYELATNERS

TPPALSTRMGLIAAVLIGQIAIEVGLSEPEYIIVSLAATETPTTPEYELATNERS

CPHDTKCFISTYKRAHRLYSANVMASIKSLQTPYLMPLIPNGKALMVOYLVRTAKPG

AKVRPSIVHPNRRLROPTNS"

1158..1164

/gene="spovaf"

/citation=11

1245..1250

/gene="lysa"

/note="putative"

1245..2664

/gene="lysa"

/note="putative"

1345..2664

/gene="lysa"

/codon\_start=1

/transl\_table=11

/product="diaminopimelate decarboxylase"

/protein\_id="AA67473.1"

/db\_xref="GI:410117"

/translation="MPLHGTSRONHGLIEIGVDALYLAEKTYGTPLYVYVALIRER

AKSPKAPFISGLKAOYAVASAKFSPVAMTOLAEDEGSLIDVVGSGELYAVAAAGCPA

ERHIFHGNNKSRERLRALEHRICIVYDNFTALLEDLCKETGHSIDVLRITTPGV



```

conflict
1610. .1616
/gene="lysa"
/citation=[1]
1751. .1757
/gene="lysa"
/citation=[1]
2496. .2502
/gene="lysa"
/citation=[1]
2504. .2510
/gene="lysa"
/citation=[1]
2666. .2694
complement(2666. .2694)
/note="ORFx19"
/note="ORFx19"
/codon_start=1
/transl_table=1
/protein_id="AA67474.1"
/db_xref="GI:410118"
/translacion="MKKIMIGMLAAVALLMWPKVSILDAADVIVITLADLSBESDK
QVLEDMNVDPNATVTVTNKEEHEYLGRKYSNMOIGSRAISSITITAKSGSLESDPK
THNIGSIDEMVLTALMTAGYKDAVVYVTAPEVSGATLGLIKAEVSEDAISDPD
VQVANAQELVYTTSELGDKICGNPMAALIAKIKKEPAKGVDPNKADILEKOVDAASDL
NVTLLDSQKNOLVSLFNKMKKADIDMGVSDQLKAKDKITKFLIESDGKNFLQKVIDI
FVYSIWMNAIVSIFK"
3347. .3353
/citation=[1]
3413. .3419
/citation=[1]
3428. .3434
/citation=[1]
3503. .3509
/citation=[1]
3819. .4290
/gene="pp1b"
3819. .4250
/gene="pp1b"
3819. .4250
/gene="pp1b"
/codon_start=1
/transl_table=1
/product="peptidyl-prolyl isomerase"
/protein_id="AA67475.1"
/db_xref="GI:410119"
/translacion="MKTGVFLLEDGNKIEFELYPEAAGTVANFEKLANEGFYDGLTFE
HIVIGFVSGGCPHGCTGTGGGYIKCTEGSNPTHEAGLSMAHNGKDTGSGQFITI
VHEPQPHNGVITVYGVKYSGLIEFAKNMSNDYKKEVYEG"
3982. .3988
/gene="pp1b"
/citation=[1]
4261. .4290
/gene="pp1b"
4287. .4320
/note="inverted repeat"
5264. .5467
/note="ORFx1"
/codon_start=1
/transl_table=1
/protein_id="AA67476.1"
/db_xref="GI:410120"
/translacion="MOVLITGLLEQYGAVLFIWMLBLALLGPGEFAGVLPQEKLN
WVLSMIAAGLIGSCGEHVILDRV"
5445. .5834
/note="ORFx2"
/codon_start=1
/transl_table=1
/protein_id="AA67477.1"
/db_xref="GI:410121"

```

Query Match	Best Local Similarity	Score	11.7%	120.4	DB: 2	Length	28206
Matches	317: Conservative	0: Mismatches	296	Indels	9	Gaps	1
QY 1	ATGCATTATGCGATTCACCTTGCACAAATATGCTTACACAGGCTTATATCCACCC	60					
DB 8281	ATGAAGCTGCGCTTGATGATCTTGCAGACGAGCGGACGACGACGATCCATCCATCCGCTC	8340					
QY 61	GTTGGCGCTGTTGATGATTAATGAAAGTGAAGTGTGTTGATTTGTTGTCACACTTGAGAAA	120					
DB 8341	GTCGGCGCTGTTGTCGTAAGAGCGACGACAAATTTGCGGAATGGCGCCCATTTAAATAT	8400					
QY 121	GCTGACAAAGCATGCGGAGGTTCAAGCACTTATATGTCACACAATAATGCTGAAGTGGC	180					
DB 8401	GGTGAAGCTCATGCAAGAGTTTCATGCTTCATATGCTGAGACACATGCAAGAGGTGCC	8460					
QY 181	ACGATTATTTACCTTGATGACCATGTATGTCATTTGCTTCAACACACCCCTGTATAC	240					
DB 8461	GACATTTTCGTTACCTGCAACCGTGCAGCCATTATGCAAAAACACGCCCATGTGCGAA	8520					
QY 241	AAATTTATTTGATTTGTAGATGACAAAAGTAGTATATACGCACAAAAG	291					
DB 8521	TTGATTATCAACTCGTGGATCAAAAAGAGTGTTCGTTGGCGATGAGAGATCCCTAATCCGCTT	8580					

-35_signal	/translation="MSYWCIVRLCPAAIIOAFASRCLLAVIYLTFPMVIFRCODKDI ELSLFPLVIGGELYEGCIRLRHLPVAPLESHVLIFFPSEQTLITAFVIFGCAY LFNHSKSHRLQTVIAFPCLLSFYSSSG"
gene	6022..6027 /gene="sips" /note="putative"
-10_signal	6022..7323 /gene="sips" 6045..6051 /gene="sips" /note="putative"
-35_signal	6683..6688 /gene="sips" /note="putative"
-10_signal	6705..6710 /gene="sips" /note="putative"
CDS	6740..7294 /gene="sips" /codon_start=1 /transl_table=1 /product="signal peptidase" /protein_id="AA067478.1" /db_xref="GI:410123"
terminator	/translation="MKSENVSKKSKILEMAKAIVAVIALIRNEIPARYVVDGDM YPILHNREVFVNMTKYIGEFDGDIVLVINGDVIYVKRIICLPGVTEMKNDOLYI NGKWDPEPYLAANKRAKODGFDLTDFDPVKVPDNRKVFMGDNRNMSDSNSGL ETKQIQGISKFVEYPIPNEMRKTN"
-35_signal	7299..7323 /gene="sips" 7388..7393 /note="putative; ORFX4"
-10_signal	7410..7416 /note="putative; ORFX4"
CDS	7529..7873 /note="ORFX4" /codon_start=1 /transl_table=1 /protein_id="AA067479.1" /db_xref="GI:410123"
terminator	/translation="MGRIKRTITLLVLLLLLAGTWYINDIEKDVPIAIGTLSSE EEETIOEYVTKIDGSEYHGVAENGTIIENGKKINMODISDIKEDKIKIAPFSKSR IDGLIKAKVND"
-35_signal	7872..7911 /note="ORFX4"
-10_signal	7937..7942 7960..7965 8126..8137 /note="putative"
Query Match	11.7%; Score 120.4; DB 2; Length 28206;
Best Local Similarity	51.0%; Pred. No. 7.5e-13;
Matches 317; Conservative	0; Mismatches 296; Indels 9; Gaps 1;

  

Db	1	ATGCATTATCGCATTCACCTGCCAAATATPGGTACAGCTCAACAGGGTTATATCCACC	60
Db	8281	ATGAAGCTGCGCCTTGATGACTTTCGCAGACGAGGCGCAAGACACGAATCCATTCGGCTC	8340
Query	61	GTTGGCGCTGTTGTAGTTAATGAAAGTGAAGTTGTTGGTATTTGGTGCACACTTGAGAAA	120
Db	8341	GTGCGCGCTGTGTCGTAAAGACGACGACAAMTTGCGGAATGGCGCCCATTTAAATAAT	8400
Query	121	GCTGCACAGCATCGGAGGTTTCACGACTTGATATGCGCACACAATAATGCTGAAAGTGGC	180
Db	8401	GGTGAAGCTCATGCAAGAAGTTCATGCCCATATGCTGAGAGCATGCGAGAGGTGCC	8460
Query	181	ACGATTATTTACCTTAGCTGACCATGTATGTCATTTGGTTCACACACCCCTGTATTAC	240
Db	8461	GACATTTCGTTACCTCGAACACCGTCAGCCATTATGCGAAAAACACGCCCATGTGCGAAA	8520
Query	241	AAATATTATGATTGTAGATGACAAAAGTAGTATATACGCACAAAAG-----CAAT	291
Db	8521	TTGATTATCAACTCGTGATCAAAGAGTGTTCGGGCGATGAGAGATGCTTAATCCGGCTT	8580

OY 292 TCGTTGACACACATGTCATGACGCTTACGGCTCACGCTATTGACGTTGAATGCCCTT 351  
 DB 8581 GTGGCTGGAAGAGGATGATGACATGATGAAGAAGCTGGCTTATGAGTAAGGAAAGCATC 8640  
 OY 352 GATGATGAACGGGCGATCACAATATATACGACGCTTTTAAACCAAAAGCAAGCAACTG 411  
 DB 8641 CTGGCAGACAGCGCGAGAGAGCGCTGAATGAAAAATTTCTGACACTTTTATGAGGACAGCCCTT 8700  
 OY 412 CCACAATAATACACTGAAAGTATCTGCAAGTTTATGATGTTAAACAGCAATGATTAATGA 471  
 DB 8701 CCGTACGCTGATTAAGAGCGGCTGACGCGCAAGATGATACACAGCGGCT 8760  
 OY 472 CAAAGTCATGATTAATCAACAAGAGTTAAACCAAGATGCTATTAAGTTAAGCATCGA 531  
 DB 8761 GACAGCAAAATGATCAGCTAGAGGCTGCAAGACAGAGATGCTCAGCAATACAGGAAACA 8820  
 OY 532 CACGACGACGAGTGTAACTGCAAGACGTTGCAATGATGATGATCAATATACGATAC 591  
 DB 8821 CACCAAGACATTTTACTGCGAGTTGGCAGCTGATGAAGCCGACATCGAGCTTAACCTGC 8880  
 OY 592 CGTATTCAGATGGAAGAAACC 613  
 DB 8881 AGACTGCCGAATGTACAAAC 8902

RESULT 10  
 BSUB0013/C DNA BCT 26-NOV-1997  
 LOCUS  
 DEFINITION Bacillus subtilis complete genome (section 13 of 21): from 2395261 to 2613730  
 ACCESSION 299116 AL009126  
 VERSION 299116.1 GI:2634723  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Bacillus subtilis.  
 Bacillus subtilis  
 Bacteria: Firmicutes: Bacillus/Clostridium group; Bacillus.  
 1 (bases 1 to 218470)

Kunst, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G., Azevedo, V., Bortner, M.G., Bessières, P., Bolotin, A., Borchert, S., Boriss, R., Boursier, L., Brans, A., Braun, M., Brignelli, S.C., Bron, S., Brouillet, S., Bruschi, C.V., Caldwell, B., Capuano, V., Carier, N.M., Choi, S.K., Codani, J.J., Conteron, J.F., Cummings, N.J., Daniel, R.A., Denizot, F., Devine, K.M., Dusterhoft, A., Ehrlich, S.D., Emerson, P.T., Entian, K.D., Erlington, J., Fabret, C., Ferrari, E., Foulger, D., Fritz, C., Fujita, M., Fujita, Y., Fuma, S., Galizzi, A., Galleron, N., Gim, S.Y., Glaser, P., Goffeau, A., Golightly, E.J., Grandi, G., Guisepi, G., Guy, B.J., Haga, K., Halech, J., Harwood, C.R., Henaut, A., Hilbert, H., Holst, S., Hosono, S., Hullo, M.F., Itaya, M., Jones, L., Joris, B., Karamata, D., Kasahara, Y., Klauer, Blanchard, M., Klein, C., Kobayashi, Y., Koeller, P., Konings, G., Krogh, S., Kumano, M., Kurita, K., Lapidus, A., Lardinois, S., Lauber, J., Lazarevic, V., Lee, S.M., Levine, A., Liu, H., Masuda, S., Mauch, C., Medigue, C., Medina, N., Mellado, R.P., Mizuno, M., Moestl, D., Nakai, S., Noback, M., Noone, D., O'Reilly, M., Ogawa, K., Ogiwara, A., Oudega, B., Park, S.H., Parro, V., Pohl, T.M., Portelle, D., Porwolik, S., Prescott, A.M., Prescan, E., Pujic, P., Rivelle, B., Rapoport, G., Rey, M., Reynolds, S., Rieger, M., Rivalet, C., Rocha, E., Roche, B., Rose, M., Sadale, Y., Sato, T., Scanlan, E., Schleich, S., Schroeter, R., Scoffone, F., Sekiguchi, J., Sekowska, A., Seror, S.J., Serron, P., Shin, B.S., Solio, B., Sorokin, A., Tacconi, E., Takagi, T., Takahashi, H., Takemaru, K., Takeuchi, M., Yamakoshi, A., Tanaka, T., Terpstra, P., Tognoni, A., Tosato, V., Uchiyama, S., Vandenbol, M., Vannier, F., Vassarelli, A., Viari, A., Wambuit, R., Wedler, E., Wedler, H., Weitzner, T., Winters, P., Wipat, A., Yamamoto, H., Yamane, K., Yasumoto, K., Yata, K., Yoshida, K., Yoshikawa, H.F., Zumbstein, E., Yoshikawa, H. and Danchin, A.  
 TITLE The complete genome sequence of the gram-positive bacterium Bacillus subtilis  
 JOURNAL Nature 390 (6657), 249-256 (1997)  
 MEDLINE 98044033

REFERENCE 2 (bases 1 to 218470)  
 AUTHORS Kunst, F., Ogasawara, N., Yoshikawa, H. and Danchin, A.  
 TITLE Direct Submission  
 JOURNAL Submitted (18-NOV-1997) I. Moszer, A. Danchin, Institut Pasteur, Regulation de l'Expression Genetique, 28 rue du Docteur Roux, 75724 Paris Cedex 15, FRANCE. E-mail: moszer@pasteur.fr, adanchin@pasteur.fr. Phone: +33 (0)1 45 68 84 41, Fax: +33 (0)1 45 68 89 48  
 FEATURES  
 source  
 Location/Qualifiers  
 1. 218470  
 /organism="Bacillus subtilis"  
 /strain="168"  
 /db\_xref="taxon:1423"  
 complement(7. .18)  
 /gene="cmk"  
 complement(7. .696)  
 /gene="cmk"  
 complement(22. .696)  
 /gene="cmk"  
 /EC\_number="2.7.4.14"  
 /note="alternate gene name: jofC, ypfC"  
 /codon\_start=1  
 /product="cytidylate kinase"  
 /transl\_table=11  
 /protein\_id="CAB14221.1"  
 /db\_xref="GI:2634724"  
 /db\_xref="SWISS-PROT:P38493"  
 /translation="MEKLSIADIDCPAAGKSTVAKIYAEKSYIYIDTGAIRATY  
 ALOENVDLDEKLAELKRTDILITKDGQKVFVNGDVTETAIPTDEISNOVIA  
 AKHSYVREMYKROODIGKGVMDGRDIGHVLPNMEVKIFLLASVEERKRYEE  
 NVKSPDYVETLIEIARDKLDSEKSPILRKEDALEIDTSLTSLQEVADKLLA  
 VQKSR"  
 complement(775. .951)  
 /gene="ypfB"  
 complement(775. .951)  
 /gene="ypfB"  
 /function="unknown"  
 /note="alternate gene name: jofB"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="CAB14222.1"  
 /db\_xref="GI:2634725"  
 /db\_xref="SWISS-PROT:P38492"  
 /translation="MKTFRLLKILFTIATILLCVQLFPHYQHTLPYVKIYQYEGV  
 DKMEENNRLETKH"  
 complement(996. .1649)  
 /gene="ypfA"  
 complement(996. .1649)  
 /gene="ypfA"  
 /function="unknown"  
 /note="alternate gene name: jofA"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="CAB14223.1"  
 /db\_xref="GI:2634726"  
 /db\_xref="SWISS-PROT:P38491"  
 /translation="MIEGENVLLYEINELKKAKSAVSVIENNELLAIVPDVYTG  
 RTVILHNDVEVEVEGKDEYVRFISRKGVKXKRLQMICEMPPREKMKIQROY  
 VYTDVAVDVOIQGNHEEIRLTYSINISAGIIVVADISFGESGLRIITLPEEH  
 TRQIFEAIVRIFNDPSEKRRKMLFVSEIATAGQOALLOYCIRQNKRRKAME"  
 complement(1711. .3094)  
 /gene="ypb"  
 complement(1711. .3094)  
 /gene="ypb"  
 complement(1742. .3094)  
 /gene="ypb"  
 /function="unknown"  
 /note="alternate gene name: jofB, yzua  
 similar to hypothetical proteins"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="CAB14224.1"  
 /db\_xref="GI:2634727"





```

/translation="MKVFTGATGCVGRHIVRELVRLNGEYVHAGVRLSKLEKFGNO
VKGIYVDEKDISIREALGVNDPFIHLGILYERKKGITERHAYGHTKMLVEVS
KGVNKKFLEMLSGHIDEAPSRVHQTKRAAEVINSGLNTYIFERSIILGEOKLF
EDMYKITKYIPVALLPDPFGVQOPADVRAKAYEALNPEDRIYLCSTKVTY
FKELADIESYMDRKVLMIPMKRYEAFALYKIEIPEPSSDQIKMMKKNVGV
LKDEATSNGLKICKREPIREESLKMSLENFIIIGK"
6098..7165
/gene="cpx"
/note="aq_136"
6098..7165
/gene="cpx"
/codon_start=1
/translation=11
/product="cytochrome c peroxidase"
/protein_id="AAC06485.1"
/db_xref="GI:2982865"
/translation="MRKGLIMAVPAGFVVAKEKIDKELKMAROYFKPLPVANP
ONPTEPKYKGLMYYDRPLSKGLISCTCINLARYCDINLPTSGHWAIGPRRA
PVIYNAIHLAOFWDGRADVEQALGPVNPLEMATERNAYKTLKSIDYEVLKFK
AFPNEDPVYKNIKAIGAFERTLMTPSRDEFLGNTKALTEOERKGLKEPIEVOC
VACHNPGVGVGNMFAKGMTEYKMTYVYLVGKPAIKVDGFRGVTKKEEDMEYK
VPSLRNIENHYPYEHDSVWSLEDVAVRIMAKTQLOKELTDQVKDIYAFLLKALTKIP
KHALEVPBLPSPDKTPKPEL"
complement(7162..8232)
/gene="ribD1"
/note="aq_138"
complement(7162..8232)
/gene="ribD1"
/codon_start=1
/translation=11
/product="riboflavin specific deaminase"
/protein_id="AAC06487.1"
/db_xref="GI:2982867"
/translation="MIREIDKNYMLALSLAKRKGTTHNPVGAIVVYEGKIVIG
YHEKAKGPHAEVMAIGAGEKAGATLYLEPCTHGRPTDARISGIRVYA
TIDPNPLMSCKVEKLRNAGIEVDVCEAELENDFTTITQEPHITLMAQTL
DGKLTALVSSKMTSKESKRVAILRRENTAVLVGNVTKIDPHITVRFVLEOP
VRILDPLEVLPSARVLANTEAPTVITKKEKKEKLEKGVOLIKGFMLKNIL
KKLELEIMHLEWEGSPRLTSLKEGFDRIYVFIAPKIMGESLIGDGIRESIES
LKVRKKVENLGGDLVIFEKRY"
complement(8229..9173)
/gene="ribP"
Query Match 11.3%; Score 116.4; DB 1; Length 14364;
Best Local Similarity 50.8%; Pred. No. 4.2e-12;
Matches 386; Conservative 0; Mismatches 361; Indels 13; Gaps 4;

```

```

Db 7845 TCGCAGAGAGAGCCAGAGAACTCAGCAGAGATTTTACTTACTACTCAGAAAGA 7786
Oy 412 CCACAAATTACAGTAAGTATCTGCAAGTTAGATGTAACAAGCAATGATATGA 471
Db 7785 CCTTACATTAACGTAATAATGGCCAGACCTGACGGAACCACTCGGACCTCAGCGGA 7726
Oy 472 CAAGTCATGCTTACTTACCAAGAGGTTAAACAAGATGCTTAAAGTAAAGACATCA 531
Db 7725 AGCAGTAAGTGATTAACCTCAAGGAAAGCAGAAAGTAGCTCATCTTAAAGAGAG 7666
Oy 532 CACAGCAGTGTACTGGAAGACGTAAGTAAATAGTATGATCCAGATTAATCA 591
Db 7665 GCAACTGGGTTCTGTAAGGTGTAAACACAGTAAAGAGCAGCCACACTTAACGTC 7606
Oy 592 CG--TATTCAGATGAGAAACCCCTATAAGATATATATGCTTAAGTGGGATATT 648
Db 7605 AGTTTGTTCACAGGAAACCAACCTGTAAGTATTTCTGACCTGAACCTGAAGTT 7546
Oy 649 CATTTTAATCAGCAAAATTTATCAAGATCAACACCAATTTGATATATATCAAAAT 708
Db 7545 CCACTGAGTGTAAAGTTCTGAATACGGAAGACCTCCAAATTTGT-TATTAAGAAAA 7487
Oy 709 CCAATTTTAACACCAATCAACACATTTTGAATTTATTT 748
Db 7486 GGAGAAATGAAAAAGCTGAAAAACCTCAAGAAAAAGAGAGTT 7447

RESULT 12
BARIGENS 4286 bp DNA BCT -07-AUG-1997
LOCUS B. amylioliquefaciens ribB, ribG, ribA, ribH & ribT genes.
DEFINITION X95955
ACCESSION X95955
VERSION X95955.1 GI:1592687
KEYWORDS deaminase; GTP cyclohydrolase II; lumazine synthase; reductase;
ribA gene; ribB gene; ribG gene; ribH gene; riboflavin biosynthesis
operator; Riboflavin synthase; ribT gene.
SOURCE Bacillus amylioliquefaciens.
ORGANISM Bacillus amylioliquefaciens
Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillaceae;
Bacillus.
REFERENCE 1 (bases 1 to 4286)
AUTHORS Gusarov,I.I., Krenova,R.A., Podchernyaev,D.A., Iomantas,Y.V.,
Abalakhina,E.G., Stoinova,N.V., Perumov,D.A. and Kozlov,Y.I.
TITLE Riboflavin biosynthesis genes of Bacillus amylioliquefaciens:
primary structure, arrangement and regulation
JOURNAL Mol. Biol. 31, 446-453 (1997)
REFERENCE 2 (bases 1 to 4286)
AUTHORS Gusarov,I.
TITLE Direct Submission
COMMENT Submitted (27-FEB-1996) I. Gusarov, GNIgenetika, 1st Dorozhny pr.
1b, 113545, Moscow, RUSSIA
On Oct 4, 1996 this sequence version replaced gi:1212772.
FEATURES
source
location/Qualifiers
1..4286
/organism="Bacillus amylioliquefaciens"
/strain="A 50"
/db_xref="taxon:1390"
321..348
promoter
/note="riboflavin operon"
653..1768
gene
/gene="ribG"
653..1768
CDS
/gene="ribB"
/codon_start=1
/translation=11
/product="deaminase"
/protein_id="CAA65189.1"
/db_xref="GI:1592688"
/db_xref="SWISS-PROT:P70814"
/translation="MEEYVKNFATLRLARREGOTQRPVLYGVAVVKKRQVYGAHLQ
YGAHAEVHAINNAGSLAKGADLYTDEPCHSKTPPCAEILMKSGIKRVFTAVEDP
NPLVAGKGIIMLEAGIEVYTGILLRODAEELNKMFLHFMRTGLPYVTLKAAASLDKGT

```

[illegible]

	Query Match	10.9%	Score 11.2	DB 2:	Length 4286;
	Best Local Similarity	52.0%	Pred. No.	2,76-11:	
	Matches	305;	Conservative	0;	Mismatches 273; Indels 9; Gaps 2;
Oy	1 ATGCATTTATGCGATTTCGAACCTTGGCAAAATTTGTGTACAGGTCAACAAGCGTGTTAATCCACC	60			
Db	668 ATTGATTACGGCATAATGACTTGCACGACGCGGTGAAGSACACAGCACGCCGATCTCTTT	727			

QY	61	GTGGCGCTGTTGGTTGTTAATGAAAGTGGAGATGTTGGTATTTGGTGGACATCTTGAAAA	120
Db	728	GTGGCGCGCCCTGCTGTGTAAAAAAGCAGACAGATCGTCGAAATGGCGCCCACTTGCATATC	787
QY	121	GGTGCAAGCATGCGGAGGTTTCAGCCACTTGATATGGCAACAACAAATGCTGAGGTGGC	180
Db	788	GGAGAGGCCCATGCTGAGAGGTGCACCCGATCATATGGCGAGCGCATACAAAGGCGCT	847
QY	181	ACGATTTTATTTAGCTTGTAGGCATGTAATGTCATTTTGGTTCACACACACCGTGTATAC	240
Db	848	GATTGTGTATGTACCGCTTGAGCCTTGCAGCCATTACGGAAAAACACCCGCAATGTGCCGA	907
QY	241	AAATTTATTTGATTGTAGATAGCAAAAGTAGTATACGCACAAAG-----ACAATTGC	294
Db	908	CTGATCATGAATATCGGGGATATAAAAGGGTGTTATTCGCTGTGAAGATCCCAATCGGCTT	967
QY	295	TTAGACACACATGTGTGATGAGACGTTACCGGCTCACCGGTATTGAGTGTGAAT---CGCTT	351
Db	968	TGAGCGGGAAAAAGGATCACAAATGCTGGAGAGAGCGGGCATGGAAGTGAAGAAAGGGGGCTT	1027
QY	352	GATGATGAGACGGCGCATCCAAATTATACCAGACCTTTTAAAAGCAAAAAGCAAGCAACTG	411
Db	1028	CTCAGCGCAACAGCGAGAGAGCTGAATAAATGTTTCTTCACCTTCATGCGGACCGGACTT	1087
QY	412	CCACAAATTTACAGCAAGTATCTGCAAGTTTAGATGGTAAACCAAGCGGATATATATGGA	471
Db	1088	CCGTTGTGCACCTTGAAGCGCGGCCCAAGACCTTTGACGGGAAAAACAGCGCACTAAACCGGT	1147
QY	472	CAAAAGTCAAATGATTACTTAACCAAAAGAGGTTAAACAAGATGCTTATATAGTTAAGCAATCGA	531
Db	1148	GACACCAAAATGGATTATCATCAGAGGCGGCAAGGCTTGATGGCCACAGCAATACAGNAATCG	1207
QY	532	CACGACGAGTGTTAACGTGGAAGACGCTACAGTGTGAATTAGATGATCC	578
Db	1208	CATCAAGCATTTCTGTGCGAGCGAGCGACAGTCAAGCGGATTAATCC	1254

RESULT	13
SPEJ908/c	
LOCUS	SPNEU1908/16-DEC-2000
DEFINITION	Streptococcus pneumoniae clone G54, *** SEQUENCING IN PROGRESS ***,
ACCESSION	AF449930
VERSION	AL449930.1 GI:11545155
KEYWORDS	HTG; HTGS PHASE2.
SOURCE	Streptococcus pneumoniae.
ORGANISM	Streptococcus pneumoniae. Bacteria: Firmicutes; Bacillus/Clostridium group; Streptococcaceae  Streptococcus. 1 (bases 1 to 175936) Dopazo,J., Mendoza,A., Herrero,J., Caldara,F., Polissi,A., Humbert,Y., Friedl,L., Guerrier,M., Grand-Scheneck,E., Gandin,C., de Francesco,M., Buell,G., Feger,G., Garcia,E., Peltsch,M. and Garcia-Bustos,J.F. Direct Submission Submitted (31-OCT-2000) Research Department, Glaxo Wellcome, S.A., Severo Ochoa 2, 28760 Tres Cantos, SPAIN * NOTE: This is a 'working draft' sequence. * This sequence will be replaced * by the finished sequence as soon as it is available and * the accession number will be preserved. location/Qualifiers 1..175936 /organism="Streptococcus pneumoniae" /serotype="19F" /db_xref="taxon:1313" /clone="G54"
TITLE	JOURNAL
COMMENT	
FEATURES	source
BASE COUNT	52994 a 31655 c 38793 g 52447 t     47 others
ORIGIN	

Query Match	10.8%;	Score 111;	DB 84;	Length 175936;
Best Local Similarity	49.0%;	Pred. No. 3.5e-11;		



/db\_xref="GI:12723932"  
 /translation="WQEPMEKVKKKRKLIIITTSVCTUNSLVLVLCQWNLKYS  
 VYAOHLPHKVGTVNRYMAVFNNDLAIYIPNANYNVPFGVAVKTSSEGLEPGLK  
 INDIVFKKGLSEFLKPEAGGLDIDVFTPTYNKNSKGEFISTVYSEETPKTIK  
 TLOYIIVKEKINEMVGFIVSHR"  
 4681..5343  
 /gene="ykbf"  
 /note="L18811"  
 4681..5343  
 /gene="ykbf"  
 /note="EVIDENCE BY GENEMARK BIO16.01 Unknown"  
 /codon\_start=1  
 /transl\_table=11  
 /product="UNKNOWN PROTEIN"  
 /protein\_id="AAK05087.1"  
 /db\_xref="GI:12723933"  
 /translation="WVRRKIVLFVPLVILLMLTFELSTEKNLHAYSIYAOHV  
 HRTGDPVCAVIDNLDNIYIPELNKSIVPEMRVNVPSKGDITDPIQYIN  
 ELKTRQALSVGPPOYNPIGPVKYOTIYITGKAGTSFTNEFSEOCUSFDSGRIMY  
 ISNLDYKGFYNLRKKTDLTKNINEMKRDAYSIIVKKRKVPSIKLQFLYMLNKRK  
 NS"  
 5393..6004  
 /gene="ykaa"  
 /note="L19523"  
 5393..6004  
 /gene="ykaa"  
 /note="EVIDENCE BY GENEMARK BIO16.01 Unknown"  
 /codon\_start=1  
 /transl\_table=11  
 /product="UNKNOWN PROTEIN"  
 /protein\_id="AAK05088.1"  
 /db\_xref="GI:12723934"  
 /translation="WVKRRKRFLEFSLIIFVILLISGYRNLVYSIYAON  
 IPKSGTDPVCAVIFNLSSVYLPHADLSTEIDRKDNQHWQSYNRYTFYNSQDI  
 EYHRANQEVSEFSMYEFENTYLVISDKNDIYRAEFERSGRLSIDIGDSRTIPT  
 NKDEILSYMKKKLVGDIVQKKRPIINLQWLYNLNKKRFS"  
 6001..6678  
 /gene="ykc8"  
 /note="L20134"  
 6001..6678  
 /gene="ykc8"  
 /note="EVIDENCE BY GENEMARK BIO16.01 Unknown"  
 /codon\_start=1  
 /transl\_table=11  
 /product="UNKNOWN PROTEIN"  
 /protein\_id="AAK05089.1"  
 /db\_xref="GI:12723935"  
 /translation="MSWVKRRKRFKRVPLIPLIFVSVLIIGYHNNERYSIYA  
 ONIPKSGTAPVCAVIDNLDNIYIPNLEKSRVNDVNVFSDRQVLEPDEVN  
 SLICKTFEDLEGISPLNODKRTISGINNTRKKIGYVINFNTKKLLKEKVVYSNO  
 GFEKIFTEINTEKGFVSVEYKNNLATISKLOKLYSIIQKSOVPLINLOMLNWL  
 NYKRFS"  
 6694..7119  
 /gene="ykc8"  
 /note="L20815"  
 6694..7119  
 /gene="ykc8"  
 /note="EVIDENCE BY GENEMARK BIO16.01 Unknown"  
 /codon\_start=1  
 /transl\_table=11  
 /product="UNKNOWN PROTEIN"  
 /protein\_id="AAK05090.1"  
 /db\_xref="GI:12723936"  
 /translation="MMNRNTOKTDTPPKRRLPLWLIPLVGMIFLSVITLESYSKY  
 AKIYHVEKQELDEVHTINNGKQLVTSQSNFPIYIDIKASECTPIYVTLIVEN  
 DYSYVKNDMPKKIIMIEGMSAGIKLITSEHGLR"  
 7651..8739  
 /gene="ribg"  
 /note="L0163"  
 7651..8739  
 /gene="ribg"  
 /note="EVIDENCE BY HOMOLOG Y BIO02.09 Riboflavin and  
 cobalamin. INVOLVED IN RIBOFLAVIN BIOSYNTHESIS. CONVERTS

2,5-DIAMINO-6-(RIBOSYLAMINO)-4(3H)-PYRIMIDINONE  
 5'-PHOSPHATE IM70  
 5'-AMINO-6-(RIBOSYLAMINO)-2,4(1H,3H)-PYRIMIDINEDIONE  
 5'-PHOSPHATE  
 /codon\_start=1  
 /transl\_table=11  
 /product="riboflavin-specific deaminase (EC 3.5.4.-)"  
 /protein\_id="AAK05091.1"  
 /db\_xref="GI:12723937"  
 /translation="MKDEYVMDLAVLNKGGVNVNPNQVALIYKGRIGGYH  
 EKYGAHEINAFKDCNPSPEGATVLEPCAHQKQPCPEALIKKIRVYIHL  
 DPNLVSGKGIKAMREACIEVSANVLECEENKIFPYYSQGLPYMMKRYMPLD  
 KATAGOSKMTTGTETARQKVHODSRMALIVETILDDPOLSVLEGNKPIRI  
 ICDTHLRSLPASKIVSTAREIPIIATAESPEKQAPFLKAGCEIILVSCLEGRLNK  
 ELTKRIGEGIDSIILEGGTLNALSALAGVQVQYIAIKIFGKREAPSPVEHGII  
 EOPNOAVYLEGQIYLTGEDILLESLL"  
 8739..9389  
 /gene="ribb"  
 /note="L0164"  
 8739..9389  
 /gene="ribb"  
 /note="EVIDENCE BY HOMOLOG Y BIO02.09 Riboflavin and  
 cobalamin. FINAL STEPS OF RIBOFLAVIN SYNTHESIS"  
 /codon\_start=1  
 /transl\_table=11  
 /product="riboflavin synthase alpha chain (EC 2.5.1.9)"  
 /protein\_id="AAK05092.1"  
 /db\_xref="GI:12723938"  
 /translation="MTGIIIEIGEVSRLOKREKFSIRASLIEMKIGDSTN  
 GICLVTFSEFNOKEPTADIMPESLNSNIGTLNVGKAVLEAMPINFGCHYISGHI  
 DVGQILKVRKEBNAIWEIETSLMLKRYTEKSGITIDGISTVIAOVKSSFSYSII  
 PHTKETIMADKKQVDVNLNDILGKRIEKLWGDSSSMKESSEVTKERLNNGL"  
 9398..10594  
 /gene="ribb"  
 /note="L0165"  
 9398..10594  
 /gene="ribb"  
 /note="EVIDENCE BY HOMOLOG Y BIO02.09 Riboflavin and  
 cobalamin. RIBOFLAVIN SYNTHESIS"  
 /codon\_start=1  
 /transl\_table=11  
 /product="GTP cyclohydrolase II / 3,4-dihydroxy-2-butanone

Query Match 10.3%; Score 105.8; DB 1; Length 11167;  
 Best Local Similarity 49.5%; Pred. No. 3,7e-10;  
 Matches 334; Conservative 0; Mismatches 332; Indels 9; Gaps 2;

QY	1	ATGAGTTATGCCATTCACCTGCAATATGATGACAAAGTCAACAGCTGTATATCCACC	60
DB	7672	ATGGATTATGCAATTGTACTTGCACAAAGAGAGGTGTAAATGTCAACCTTAACCTCA	7731
QY	61	GTGGCGCTGTGTAGTTATGATGAGTAGATGTTGTATGTCACACTTGAGAAA	120
DB	7732	GTAGAGGCCCTTATGTCTCAAGAGGAGCAATATTGTGTCAGGTTATCATGAAAAATAT	7791
QY	121	GGTGACAGACATCGGAGGTTCAAGCAGCTGATATGGCACAACAAATCTGAAGTGGC	180
DB	7792	GGGAAAGCTACAGCGGAATCAATCTTTAAGATTTGAAGCAATCTCCAGAGGACG	7851
QY	181	ACGATTATATTAAGCTTACAGCATGCTACTCATTTTGTTCACCAACCCCGTGTATAC	240
DB	7852	ACGCTTATGTGACCTTAAGCCCTGTGCGCANTCAGGAAACACCCCTGTGTTTGA	7911
QY	241	AAATATTGATTTGATTAAGTAGCAAAAGATATACGCAACAAAGAC-----AATTGC	294
DB	7912	GGGATTATATAAATAGATTTAAGCTGTGTTATTCGGCATTTAGACCCCAATCCCTTG	7971
QY	295	TTAGACACACATGTGATGAGAGCTTACGGCTCAGCGTATTTGAGGTGATCGTGTAT	354
DB	7972	GTTCAGAGAAAGATTTAAGCAATGAGAAGCTGGAATGAGTTCTCTCAATGTT	8031
QY	355	GATGAAGGGGCAAT---CAATTTATACCAAGACTTTTAAAGCAACCAAGACACTG	411





SREALELRDGRKSRFLGKGVKAAVAANNELIAOAVGKDATNOAETIDIMLDCTEN  
 KSNFGANALIAVLANAKAAASGLPLVAYIAELNGTAGVSMPLPMNINIGEHA  
 DNNVDIOEMIQPVAGKTLREALRIGAEVFNHIALVAKGSTAVGEGGAPNINAS  
 NADALACEIKAEVAGKGYLGKGVKAVLAMPDASSEFNENKEMKGEKSGOEETH  
 YLEELCKEPIVSIIDGODESDMEGVAOTKYLGRVOLVGDLLFTVNTKILKEGIEK  
 GIANSILIKFNQISLETETIAIKMAKADGATVAVSHSGEEDAT IADLANGTAAG  
 IKTGSMNSDRITAKINOLIRIEALEAFRAGTAPAPGLAKAVKGA"  
 complement(3855. .5060)  
 /gene="H10933"  
 complement(3855. .5060)  
 /note="similar to SP:P37631 PID:912464 GB:U00096  
 PID:2367234 percent identity: 58.35; identified by  
 sequence similarity; putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="conserved hypothetical transmembrane protein"  
 /protein\_id="AAC22591.1"  
 /db\_xref="GI:1573954"  
 /translation="MSQYSENIIGAGAGLFCQAQLAKLKSQVTFPDNGKIKRKL  
 MSGGFCNFTNLEVPAYHLSQPHFVKSSALARYNMFISLVAEGITVHKEIGOL  
 FCDGAEQIVEMLKSECDKYGAKILRESEVERIODEKRVLYQVNSTOMCKNL  
 IVATGGLSPGAGTPEGQIAEFGPIPIPRASLVPEYETDPLTALSGISLPV  
 TTTALCGSEVNOILLFTRHGISGPAVLQISNWDQTESVEIDLPHNHVEEINAKO  
 SSPKMLKTIIVRLPKKIVELMIEQIVQDEYVIANISKVRKKNLVDFLIHMEFTNG  
 TEGYRTAEVTMGVDYFKVYISSKTMESNOVSGLYTIGVLDVYGLGYNFWQAMSSAY  
 ACALSISSRO"  
 complement(5057. .6211)  
 /gene="H10934"  
 complement(5057. .6211)  
 /note="similar to SP:P32711 GB:X72298 PID:396410  
 PID:404307 GB:U00096 percent identity: 50.41; identified  
 by sequence similarity; putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="cytochrome C-type biogenesis protein (nrf)"  
 /protein\_id="AAC22592.1"  
 /db\_xref="GI:1573955"  
 /translation="MKSILDIITRGFIQVOKTVEFATALLFAFLFTQAEVNTYQFO  
 NODDTRAVELAKSLRPOCONONVESPAYLQRELVYVMEVQEGSNQOIIDKMT  
 ARFGDEVNKKPEKMTALIMLIPVALLIAVLIFSNRRKQSEKVAADROENDEI  
 ISLPTRSSSPRKQSEPKSLSGKYNKITYPLFTLLIAPITYFSLDRSRVQGE  
 QSMITQHNQVBNDEHNENYIEKLNKMLRTDPNNAETWLOLGEAVYVONNEFSDALV  
 AYSNAEKLSGSKPNILGLAATLAAVYQAGQOMQKVEQLNEALADKNEVSSLAT  
 IALENRQYQOAGMYLQOLIDSGNAAVDRSRVYQRMKMDLQREGKGNP"  
 complement(6208. .6738)  
 /gene="H10935"  
 complement(6208. .6738)  
 /note="similar to SP:P33926 PID:405920 GB:U00096  
 PID:1788523 percent identity: 42.60; identified by  
 sequence similarity; putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="thiol:disulfide interchange protein (dsb)"  
 /protein\_id="AAC22593.1"  
 /db\_xref="GI:1573956"  
 /translation="MNRKLIETPLFVLGVCILLAGLNDPKRTIASALIDKPYEE  
 YOANLHPSOAIYSPKEPKPFLNVMGSMGVCYKEEHLIETIKTLPIVGVDRDN  
 RONGIKMLKRMGNPFIITIDSHGELGLKGVGADAPETVYVNDENVIRRHSGLLDKE  
 TWQYVFLPKIEFLANK"  
 complement(6738. .8645)  
 /gene="H10936"  
 complement(6738. .8645)  
 /note="similar to SP:P32710 GB:X72298 PID:581146 GB:U00096  
 PID:1790511 percent identity: 42.42; identified by  
 sequence similarity; putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="cytochrome C-type biogenesis"  
 /protein\_id="AAC22594.1"

/db\_xref="GI:1573957"  
 /translation="MLPELGEFLLLATASAFLLAVPQGLFKKNPILLINAMPISY  
 IETLATTSIGLAFSAFVADFTLEYVAHNSOLPTEFKYAAWTGEGSMLFVLS  
 LSLMAFALEFRKNDREFSOSLSLILICGFVFLFIPNSPGRIFPPAPBGRDL  
 NPMIDVGLIFHPPLLIVGYGFAVNFPMASLALYNSAQIARSMGWLVSMLFL  
 TIGIVLGMWAVYELGKGGWFMVDVENASLMPILLGLALHSLMATEKQGVFTWTT  
 9.3%; Score 95.2; DB 3; Length 18073;  
 Best Local Similarity 49.8%; Pred. No. 3.1e-08;  
 Matches 300; Conservative 0; Mismatches 293; Indels 9; Gaps 2;  
 Query Match  
 QY 1 ARGGATTATGCCATTTCACTTCAATATGTCACAAAGTCACAGGTGTAATCCACC 60  
 DB 15900 ATGCAACGTCCTTAGATTGGCAGCAAGCAATATACACACCAACCCATCG 15959  
 QY 61 GTTGGCGCTGTTGTAATGTAATGTAAGTATGTTGTAATGTTGTCACATGGAANA 120  
 DB 15960 GTGGAGTGGCTATGTTGTTAAATGTTGTAATGTTGTTGTTGTTGTTGTTAAGCA 16019  
 QY 121 GGTGACACACATGCGAGGTTCAAGCATTTGATGACCAACAAATGCTGAAGTGGC 180  
 DB 16020 AGCAACCTCATGCTGACGCGTTCCTGATGCAAGTGGTGAANAATGCTAAAGAGCA 16079  
 QY 181 AGCATTTATATTAAGTTAGACCATGTAATGTTGTTGTTCAACACCACTGTTTAAAC 240  
 DB 16080 ACCGCTTACGTCACTTACAGGCTTGGCTCATTTATGTCGACACACCTTGTGATTA 16139  
 QY 241 AAAATTATTGATTGTAAGTATGCAAAAGTATACGCAAAAG-----ACAATTG 294  
 DB 16140 GATTTATTTAGAGCGGCGTGTGTAAGTATGCGGCTATGCAAGATCCAAATCCTCA 16199  
 QY 295 TTAGACACACATGTTGATGATGACGCTTACGCGTCACGCTATTGAGGTT--GAATGCGTT 351  
 DB 16200 GTGCGAGAAAGGTTTGAANAATGTTGCTGATGCGGATGCAAAAGTACGGTGAATTTA 16259  
 QY 352 GATGATGAACGGGCATCAATTTATACAGACTTTTAAAGCAAAAGCAAGCAACTG 411  
 DB 16260 TTGAACGATCAAGCGGAAAAAATAAAGTTTAAAGCAAGATGCTGAAGCTATG 16319  
 QY 412 CCACAATATACAGTGAAGTATCTGCAAGTTTATAGATGTAACAAAGCAATGATATGA 471  
 DB 16320 CCTTTGTTCAACTAAACTTGCATGATGATGATGATGATGATGATGATGATGATGATG 16379  
 QY 472 CAAAGTCAATGATTAATAAAGAGGTTAAACAGATGCTATTAAGTAAAGATGCA 531  
 DB 16380 GAAAGTAAATGATTAAGGCGCCAGATGCTGTCGAGCTGCAAAAATAAGCGAGCAAAA 16439  
 QY 532 CACGACGAGTGTAACTGAGAGCTTACAGTTGATTTAGATGATCCCAATATACTACA 591  
 DB 16440 TCATCCGCACTTTTATCTTCCACACAGTCATTGTCAGATGATCCAGCTTAATGTA 16499  
 QY 592 CG 593  
 DB 16500 CG 16501

Search completed: May 7, 2001, 12:17:16  
 Job time: 6132 sec



XX Rosenberg M, Trautl CM, Ward JM, Warren RL;  
 DR WPi: 1999-347572/29.  
 DR P-PSDB; Y15913.  
 XX  
 PT New Staphylococcus aureus riboglycopeptide and polynucleotide useful in  
 PT the treatment of gastric ulcer and gastritis  
 XX  
 PS Claim 6; Page 5; 48pp; English.  
 XX  
 CC The present sequence encodes a Staphylococcus aureus pyrimidine  
 CC deaminase and pyrimidine reductase (riboglycopeptide). Staphylococcus  
 CC aureus riboglycopeptide and its antagonists are used to treat individuals in  
 CC need of them. Disease related to expression or activity of riboglycopeptide  
 CC can be determined by analysing the nucleic acid sequence encoding riboglycopeptide  
 CC or detecting the riboglycopeptide in a sample. riboglycopeptide can also be used  
 CC to identify antagonists or agonists. riboglycopeptide, or its related nucleic acid,  
 CC also has use as a vaccine to induce an immunological response in an  
 CC animal. Antimicrobial compounds (e.g. agonists and antagonists of riboglycopeptide)  
 CC especially broad-spectrum antibiotics, may be of use in the treatment  
 CC of Helicobacter pylori infection. This should decrease the advent of  
 CC H. pylori-induced cancers, such as gastrointestinal carcinoma. The  
 CC treatment should also cure gastric ulcers and gastritis.  
 XX  
 X0 Sequence 1029 BP; 376 A; 162 C; 198 G; 293 T; 0 other;

Query Match	100.0%;	Score 1029;	DB 20;	Length 1029;
Best Local Similarity	100.0%;	Pred. No. 5.6e-230;		
Matches 1029;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

OY	1	ATGATTAATGCGATTCAACTGCAAAATATGATACAAAGTCMAAACGGTGTAAATCCACC	60
Db	1	atgattatgcatcactcaacttgcgaatcatgttacaagtgcaaacagtglttaatccacc	60
OY	61	GTTGGCGCTGTTGTAGTTAATGAANGTAGAGTTGTTGTATGGTCACACTTGAGAAA	120
Db	61	gttggcctgttctagttaataagaagtagatgttgttacttggtgcacacttgaagaaa	120
OY	121	GGTGACAAAGCATGTGGGAGGTTTCAAGCACTTGATGTGGCACACAAAATGCTGGAAGGTGGC	180
Db	121	gtgtacaagcatgtggagaggttccaagccttgatcttggcacaacaaatgtctgaagtgcg	180
OY	181	ACGATTATATTTACGTTAGAGCCATGTAGTCATTTTGGTTCAACACCACCCTGTTTAAC	240
Db	181	acgattatatctagcttagagccactgtagtcatttctgttccacaaccacctgtttaac	240
OY	241	AAATTTATTTGTTTGAAGATAGCAAAAAGTGTATGCGCAACAAAAGCAATTTGTTAGAC	300
Db	241	aaattatttgcctgttaagatagcaaaaagtgltatacgaacaacaaagcaatcgtttagac	300
OY	301	ACACATGGGTGATGAGACGTTACGGGGCTACAGGTAATTGAGGTTGAATTCGTTGATGAA	360
Db	301	acacatgtgtatgtagacgcttagcggctccaagttatgaagtgtagtcgttgtatgaa	360
OY	361	CGGGCATCACAAATATPCCAAAGACTTTTATTAAGCAAAAGCAAGCACTGCCCAAAAT	420
Db	361	cgggcatacacaattatpccaaagactttttaaagcaaaaagcaaacctggcacaatc	420
OY	421	ACAGTGAAGAGTATCTGCAAGTTTAGATGTAAACMACGAGTANTATATGAGCAACAGTCAA	480
Db	421	acagtgaagatctgcgaagtttagatcgtlaaacaagcgaatgataatgtgacaagtcgaa	480
OY	481	TGGATTACTAACAAGAGGTTAAACAAGATGCTCATATAGTTAAGACATGCACACGACGCA	540
Db	481	tggattactaacaagaggttaacaagaagtcctcataagttlaagacatcgcacgcgaagca	540
OY	541	GTTGTAACGTGAAAGACCTACAGTTGTAATGATGATCATCACAAATATACTACACCTATTYCAA	600
Db	541	gttgttaactgtgaagacgttacagttgaattagatgatccacatactatacagtatccaa	600
OY	601	GATGGAAAAAACCTATATAAAGTAATATTGTCTAAGCTGTGGAAATATTCATTTTAATCAG	660
Db	601	gattggaaaaaacctatataaagtaattattgtctaaagcttggaaatatctcatttaattcag	660

Db	601	gatgaaaaaacccctataaaaagtaattatgtctaaagtcctgggaatatatcattttaatcag	660
OY	661	CAAAATTATCAGATGAATCAACACCAATTGGATATATACGAAATTCAAATTTAACA	720
Db	661	caatttatcaagtgatgaatcaaacaccatttggatatatactgaaatccaatttaaca	720
OY	721	AGCATCAACACACTATTGGAATTTATTTACTTGAAGCTTTGGATTTAAACAATTCTT	780
Db	721	agcaatcaaacacatctggaattattacttgaagctcttgatttaacacaactctt	780
OY	781	CACAATTTATATAAAGAGAGATTGGAACTTTCCTAGTCGAGGACAGTCCACCACTACT	840
Db	781	cacaatttatataaaagagagatttggaaacttctgactgagcgagctccaacccactact	840
OY	841	TCAGAAATTCACATCTATTATATGATGAATTTATTTCTGATATATGCCCCGAAATTAAT	900
Db	841	tcgaattctcccatctatataatagaatattatctctatattatgccccgaattaat	900
OY	901	GGCGGATCTGGAAATTTATCAATTTTATCAACAACAATGATGTGATTAGATTACCAATGCG	960
Db	901	ggcggaatctggaattatccaatttatacaaacaaatgatagtatgagttacagagatcg	960
OY	961	AACCAATTGGAATTTGTCATTCGCGATTATTAATCAAAAGTGTAAATTAACCTTAGCA	1020
Db	961	aaccaatttgaatttcttcatccogagttatlaaaccaaaatgtaataacttaacttaca	1020
OY	1021	AAGAAGTGA	1029
Db	1021	aagaagtga	1029

RESULT	2
X59913	
ID	X59913 standard; DNA; 3336 BP.

DT 04-AUG-1999 (first entry)  
XX  
DE Riboflavin biosynthesis operon containing the ribG gene

KW pyrimidine deaminase; pyrimidine reductase; ribG; antagoanist;  
KW agonist; antimicrobial; antibiotic; Helicobacter pylori infection;  
KW H. pylori-induced cancer; gastrointestinal carcinoma; gastric ulcer,  
KW gastritis; ss.

OS Staphylococcus aureus.

PN W099926475-A1.

PD 03-JUN-1999.

PF 23-NOV-1998; 98WO-US24857.

PR 25-NOV-1997; 97US-0978456.

SMITHKLINE BEECHAM CORP.

PI Black MT, Burnham MKR, Fedon JC, Hodgson JE, Knowles DJC;

PI Rosenberg M, Tranl CM, Ward JM, Warren RL;  
 PI Rosenberg M, Tranl CM, Ward JM, Warren RL;  
 PI Rosenberg M, Tranl CM, Ward JM, Warren RL;

WPI; 1999-347572/29.

PT New Staphylococcus aureus polypeptide and polynucleotide useful in the treatment of gastric ulcer and gastritis

PS Example 2; Page 30-31; 48pp; English.

CC The present sequence represents the *Staphylococcus aureus*  
CC riboflavin biosynthesis operon containing the pyrimidine  
CC deaminase and pyrimidine reductase (ribG) gene. *Staphylococcus*  
CC *aureus* ribG and its antagonists are used to treat individuals in





Sequence 3336 BP; 1202 A; 459 C; 678 G; 997 T; 0 other:

Query Match 100.0%; Score 1029; DB 20; Length 3336;  
Best Local Similarity 100.0%; Pred. No. 7.6e-230;  
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 ATGGATTATTCGATTCACCTTGCACAAATATGATACAAAGCTGATTAATCCACCC 60
    |||||||
DB 1 atgattatcgcattcactcgtcaaatatgltacaaggtcaaacaggtgltatccaccc 60
OY 61 GTTGGCGCTGTGTAGTAAAGTATGAGATTGTGTATGGTCACACTTGACAAA 120
    |||||||
DB 61 gttggcgctgtgtagtaaagtagagtagtctgtgtatggtcaccacttgagaaa 120
OY 121 GTGACAAAGCATGGGAGTTCAGCACTGATATGGCACAAACAAATGTGAAAGTGGC 180
    |||||||
DB 121 gttgacaagaatcggaggttcacagccttgatagcacaacaacatgctgaagtcg 180
OY 181 AGCATTTATATATACGTTAGACCATGTAGTCATTTTGGTTCACACCCCTGTGTAA 240
    |||||||
DB 181 acgattatatactagtagagccatgtacgtcatttggttcaacaccctgtgtaac 240
OY 241 AAATTTATTTATTTAGATAGCAAAAGTATACCAACAAAGCAATTCGTTAGAC 300
    |||||||
DB 241 aaattatctgtgaagtagaagaagtagtacgcacaaagaacaaatctgttagac 300
OY 301 ACACATGTTGATGAGACGTTACGGCTCAGCGTATTGAGTTGAAATCGTTGATGAA 360
    |||||||
DB 301 acacaaggtagtagaaggttaccggctcacggtatcgtgaagtcgttagagtagaa 360
OY 361 CGGGCATCACAAATTTATACCAAGCTTTTAAAGCAAAACCAAGCAACGCCCAAT 420
    |||||||
DB 361 cgggcatcacaaattatataccaagctttttaaagcaaaagcaaacgcgtccaaatt 420
OY 421 ACAGTGAAGATATCTGCAAGTATTAGATGTTAAACAGCGAATGATTAATGCAAA 480
    |||||||
DB 421 acagtgaagatctcgaagtttagatgtaaaacagcgaatgataatgacaagtagaa 480
OY 481 TGATTACTTAACAAGAGTTAAACAAGATGTCTATTAAGTTAAGACATCGACGACGA 540
    |||||||
DB 481 tggattacttaacaagaggttaacaagaagctctataagtttaaacctcgaagaagca 540
OY 541 GTGTTACTGGAAGACGCTACAGTTGAATTAGATATCCACAAATATCTACACGATT 600
    |||||||
DB 541 ggttaactcgaagagcgtacagcttagaattagatccacaataatactacacgtatt 600
OY 601 GATGGAAGAAACCCCTATTAAGATATTTGTCTAAGCTCTGGCAATTTCAATTT 660
    |||||||
DB 601 gatggaagaaacccctataaagtaattgtctaagctctggaaatattcaatcag 660
OY 661 CAATTTATCAAGATGAATCAACCAATTTGATATATCTGAAATTCGAAATTTAACA 720
    |||||||
DB 661 caaattatcaagaatgaatcaacccaatttgaatatatctcgaataatccaaattt 720
OY 721 AGCAATCAACACATATTTGAATTTATTTACTTGAAGCTTGTGATTTTAACAACT 780
    |||||||
DB 721 agcaatcaaacacatattgaatatttacttgaagcttctgtgatttaacaacattc 780
OY 781 CACAATTTATTAAGAGAGAGTTGCACTTGTCTACTGACAGCAGCTCCACCACT 840
    |||||||
DB 781 cacaattatataaagagaggttggaacttctagtcgagcaggtlccaaacactact 840
OY 841 TCAGAATTCGATCTATTTATATAGATGAATTTATCTCATTTATGCCCCGAAT 900
    |||||||
DB 841 tcagaatcccatctctatataagatgaattatctccattatgcccgaataatt 900
OY 901 GCGGATCTGGAATATATCAATTTATCAACAATGATGATGATGATGATGATGCG 960
    |||||||
DB 901 gcggaatctggaatataatcaatttatacaacaatgagtgltggtatataccagat 960
OY 961 AACCAATTTGAATTTGTTCATTCGAGTTTATTAATCAAAATGTTAATTAAC 1020
    |||||||
DB 961 aaccaatttgaatttgttcattccgagtttatttaattcaaaatgtaatttaac 1020

```

```

DB 961 aaccaatttgaatttgcattccgagttatlaatacaaatgtlaaataacttaccga 1020
OY 1021 AAGAGTGA 1029
    |||||||
DB 1021 aagaagtga 1029

```

RESULT 5  
X60830  
ID X60830 standard; DNA; 3336 BP.

X60830;  
09-AUG-1999 (first entry)

S. aureus riboflavin biosynthesis operon.

Lumazine synthase; ribH; riboflavin; differential screening; gastritis;  
H. pylori infection; gastrointestinal carcinoma; gastric ulcer; wound;  
bacterial adhesion; cell invasion; riboflavin biosynthesis operon; ss.  
Staphylococcus aureus.

W09927128-A1.  
03-JUN-1999.

24-NOV-1998; 98MO-US25096.

25-NOV-1997; 97US-0977553.

(SMIK ) SMITHKLINE BEECHAM CORP.

Debonck C, Fedon JC, Jaworski DD, Mooney J, Palmer LM;  
Traini CM, Wang M, Warren RL, Zhong Y;

WPI; 1999-357850/30.

New Staphylococcus aureus ribH polynucleotide and polypeptide useful  
in the treatment of gastric ulcers and gastritis

Example 2; Page 34-36; 48pp; English.

The invention relates to polynucleotides and polypeptides of the lumazine  
synthase (beta -subunit) family, designated ribH. They are isolated from  
Staphylococcus aureus. ribH enzymes are involved in the synthesis of  
riboflavin. The polynucleotides and polypeptides may be employed as  
research reagents and material for the discovery of treatments and  
diagnostics for diseases, particularly human diseases. They can be used  
for diagnosis of the disease and staging of disease, and as reagents in  
differential screening methods. The polynucleotides may be used as a  
source for hybridization probes, and for screening of genetic mutations,  
serotype, and identification, and for chromosome identification. The  
polypeptides can be used to produce antibodies. The polypeptides can  
also be used in vaccine formulations, and to identify agonists and  
antagonists. These are used to prevent, inhibit or treat diseases,  
particularly of Helicobacter pylori infections, such as gastrointestinal  
carcinoma. Gastric ulcers and gastritis may also be treated. The  
polypeptides can also be used to treat wounds and in-dwelling devices to  
prevent bacterial adhesion and infection, and to block ribH protein-  
mediated mammalian cell invasion. Antagonists and agonists of ribH may  
also be employed to inhibit and treat diseases. The present sequence  
represents the S. aureus riboflavin biosynthesis operon that encodes  
genes ribG, ribB, ribA and ribH.

Sequence 3336 BP; 1202 A; 459 C; 678 G; 997 T; 0 other:

Query Match 100.0%; Score 1029; DB 20; Length 3336;  
Best Local Similarity 100.0%; Pred. No. 7.6e-230;  
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGGATTATTCGATTCACCTTGCACAAATATGATACAAAGCTGATTAATCCACCC 60



```

|||||
Db 1 atggttatgcatccaacttgcaaatatggtacaaaggtcaaacagtggttaatccacc 60
OY 61 GTTGGCGCTGTTGTTAGTTATATAGAGTAGAGTTGTTGGTATTGGTCACACTTGGAGAAA 120
Db 61 gtggtgcgtgtgtatgtatgaagtagagatgtgtgtatgtgtgtgtgtgtgtgtgtgtgt 120
OY 121 GGTGACAGCAGCAGGAGTTGACAGACTGTGATGAGCGCAACAAAGCTGAAGTGGC 180
Db 121 ggtgcaaaacagtcggaagttcaacacttgatagcacaacaagaatgctgaagtgctg 180
OY 181 AGGATTATATATACGTTAGAGCCATGTAAGTCTATTTGGTTTGAACACCAACCCGTGTAAAC 240
Db 181 agatttatattacgttagagcactgtagtcaatttgggttcaacaccaccctgtgtaac 240
OY 241 AAAATTATGATTGTATAGTATACAAAAGTATATACGACAAAGACAAATTCGTTAGAC 300
Db 241 aaattatgtatgtatgaatagcaaaagtagatatacgcaacaagaacatttcgttagac 300
OY 301 ACACATGATGATGAGACGTTACGGGCTACGGTATTTGAGTTGAAATGCGTTGATGATGAA 360
Db 301 acacattgtatgtatgaatagcaaaagtagatatacgcaacaagaacatttcgttagac 360
OY 361 CGGCGATCAACAATATACCAAGACTTTTAAAGCAAAAGCAAGCAACTGCCACAATTT 420
Db 361 cggcgatcaacaattatccaagaactttttaaagcaaaagcaagcaactgccaacaatt 420
OY 421 ACAGTGAAGAGTATTCGCAAGTTAGTATGTAACAAAGCAAGATGATATGACAAAGTCAA 480
Db 421 acagtgaaagtatctgcgaagtttagatgtaacaagcgaatgaatgtaagaagaatcaa 480
OY 481 TGGATTACTAACAAGAGAGTAAACACAGTGTCTAAAGTAAAGACATGACAGCAGCA 540
Db 481 tggattactaacaagagaggttaacaagaatgctataagttgaagatcgcagaagcaga 540
OY 541 GTGTTAACTGAAGACGTACAGTTGAATAGATGATCCAAATATATACAGTATTCAA 600
Db 541 gttgtaactggaagagcgtacagttgaattagatgctcaaatatctacacagttatcaa 600
OY 601 GATGGAAGAAACCCCTATATAAGTATTTGTTAAAGTCTGGGAATATTTATTTAATCAG 660
Db 601 gatggaagaaacccctataaagtaatatgtctaaagtcgtggaataatcttaatactag 660
OY 661 CAAATTTATCAAGATGAATCAACACCAATTTGATATATACGAAATCCAAATTTAAACA 720
Db 661 caaatttatcaagaatgaatcaacaccaatttgatatactacgaaatccaatlttaaca 720
OY 721 AGCAATCAAAACACATATTTGAATTTACTGGAAGTCTTGTGATTTAAACAATTTCTT 780
Db 721 agcaatcaaaacacatltgaatatttacttgaagtcctgtgatttaacaacaattctt 780
OY 781 CACATTTATATATAAGAGAGTGTGAAGTCTGCTAGTCTGAGAGCAGGTCACCACTACT 840
Db 781 cacaatttatataaagaagaggttgaacattgctagtcgagcaggtccaacccactact 840
OY 841 TCAGAAATTCCTCATATATATATAGATGAATTTATTTCTATTTATGCCCCGAAATTAAT 900
Db 841 tcagaaatctccatctatataatataatgaatattctctatataagcccggaataatatt 900
OY 901 GCGGATCTGCAAAATTTATCAATTTTATCAACAATGATGTGATGAGATTACAGATGCG 960
Db 901 gcggaatctgcaaaatattcaatlttatacaacaatgagtgtgaggtaccagatgctg 960
OY 961 AACCAATTTGAATTTGTTCAATTCGAGTTATATTAATCAAAAGTTAAATTAATTAATTA 1020
Db 961 aaccaatttgaatgttcatctccgagttatataatcaaaaatgcttaaaatlaactttaga 1020
OY 1021 AAGAGTGA 1029
Db 1021 aagaagtga 1029

```

```

V7453/C
ID V7453 standard; DNA: 7588 BP.
XX
AC V7453:
XX
DT 16-MAR-1999 (first entry)
XX
DE Staphylococcus aureus contig SEQ ID #142.
XX
XX Computer readable medium; vaccine; S.aureus infection; immunodetection;
XX cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
XX skin infection; surgical wound infection; scalded skin syndrome;
XX toxic shock syndrome; ds.
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FT misc-feature 661..720
FT /*tag= a
FT /note= "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"
FT
FT misc-feature 2461..2520
FT /*tag= b
FT /note= "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"
FT
FT misc-feature 4261..4320
FT /*tag= c
FT /note= "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"
FT
FT misc-feature 6061..6120
FT /*tag= d
FT /note= "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"
FT
EP786519-A2.
PD 30-JUL-1997.
XX
PF 07-JAN-1997; 97EP-0100117.
XX
PR 05-JAN-1996; 96US-0009861.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA;
PI Rosen CA;
XX
DR WPI; 1997-374922/35.
XX
PT Polynucleotide(s) and proteins derived from Staphylococcus aureus
PT stored on computer readable medium and used in the production of
PT anti-S.aureus vaccines
XX
PS Claim 1; Page 740-744; 3271pp; English.
XX
CC This sequence represents one of 5191 Staphylococcus aureus DNA sequences
CC of the invention. The DNA sequences are recorded on a computer readable
CC medium, preferably selected from a floppy or hard disk, random access
CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC the S.aureus DNA sequences allows putative functions to be assigned so
CC that protein-encoding or regulatory regions of commercial, therapeutic or
CC industrial importance can be obtained. Specifically, sequences which are
CC likely to encode antigens have been identified and these polypeptides can
CC be used in a vaccine composition against S.aureus infection. The
CC polypeptides can also be used in a kit for the immunodetection of

```



CC S. aureus in a sample. S. aureus is implicated in numerous human diseases, including cellulitis, eyelid infections, food poisoning, osteomyelitis, skin and surgical wound infections, scalded skin syndrome, toxic shock syndrome, etc. Organisms transformed with the DNA sequences can be used for recombinant production of the polypeptides. The new DNA sequences (and their fragments) are useful as primers or probes for isolating homologues of any of the S. aureus DNA sequences contained on the computer readable medium.

CC Sequence 7588 BP; 2333 A; 1341 C; 1061 G; 2613 T; 240 other;

Query Match 97.2%; Score 1000; DB 18; Length 7588;

Best Local Similarity 98.7%; Pred. No. 5.1e-223;

Matches 1019; Conservative 0; Mismatches 10; Indels 3; Gaps 1;

```

QY 1 ATGGATTATGCGATTCACCTTGCATATATGTTACAGTCCAAACAGGTGTTAATCCACC 60
DB 5849 ATGGATTATGCGATTCACCTTGCATATATGTTACAGTCCAAACAGGTGTTAATCCACC 5790
QY 61 GTTGGCCCTGTGTTAGTTAATGAAGTATGTTGTTGTTGTCACACTTGAGAAAA 120
DB 5789 GTTGGCCCTGTGTTAGTTAATGAAGTATGTTGTTGTTGTCACACTTGAGAAAA 5730
QY 121 GGTGACACAGATGCGGAGTTCAAGCACTTGATATGCGACACAAATGCTGAAGTGGC 180
DB 5729 GGTGACACAGATGCGGAGTTCAAGCACTTGATATGCGACACAAATGCTGAAGTGGC 5670
QY 181 ACGATTATATTCAGTTAGAGCCATGTATCATTTTGGTTCAACACACCCCTGTTAAC 240
DB 5669 ACGATTATATTCAGTTAGAGCCATGTATCATTTTGGTTCAACACACCCCTGTTAAC 5610
QY 241 AAAATTATGTTAGTATAGTACAAAAGTATATACCAACAAAGACAAATTCCTTTAGAC 300
DB 5609 AAAATTATGTTAGTATAGTACAAAAGTATATACCAACAAAGACAAATTCCTTTAGAC 5550
QY 301 ACACATGCTGATGAGAGCTTACGGGCTCAGGGTATGAGGTGTAATCGTTGATGATGA 360
DB 5549 ACACATGCTGATGAGAGCTTACGGGCTCAGGGTATGAGGTGTAATCGTTGATGATGA 5490
QY 361 CGGCGATCACAATTTATACCAAGACTTTTAAAGCAAAAGCAAGCCCAATTT 420
DB 5489 CGGCGATCACAATTTATACCAAGACTTTTAAAGCAAAAGCAAGCCCAATTT 5430
QY 421 ACAAGTGAAGTATCTGCAAGTTTATGATGTAACAAACGATGATTAATGACAAAGTCAA 480
DB 5429 ACAAGTGAAGTATCTGCAAGTTTATGATGTAACAAACGATGATTAATGACAAAGTCAA 5370
QY 481 TCGATTACTAACAAAGAGGTTAAACAAGATGCTATATAGTTAAGACATCGACGACGCA 540
DB 5369 TCGATTACTAACAAAGAGGTTAAACAAGATGCTATATAGTTAAGACATCGACGACGCA 5310
QY 541 GTGTTACTGGAAGAGCTACAGTTGATATGATGATCCACAATTTACTACACGATTTCAA 600
DB 5309 GTGTTACTGGAAGAGCTACAGTTGATATGATGATCCACAATTTACTACACGATTTCAA 5250
QY 601 GATGGAAGAAACCCCTATAAAGTAAATATGCTCTAGCTGGGAATATTTCAATTTAAATCAG 660
DB 5249 GATGGAAGAAACCCCTATAAAGTAAATATGCTCTAGCTGGGAATATTTAAATCAG 5190
QY 661 CAAATTTATACAGATGATCAACACCAATTTGGATATATCTGAAATTCCAAATTTTACA 720
DB 5189 CAAATTTATACAGATGATCAACACCAATTTGGATATATCTGAAATTCCAAATTTTACA 5130
QY 721 AGCAATCAACACATATATGAATATTTACTCTGGAAGTCTGTGCTTAAACAACAAATCTT 780
DB 5129 AGCAATCAACACATATATGAATATTTACTCTGGAAGTCTGTGCTTAAACAACAAATCTT 5070
QY 781 CACAATTTATATAAAGAGAGGTTGAAGTCTGCTAGCTGACGAGGCTCCAAACCACTACT 840
DB 5069 CACAATTTATATAAAGAGAGGTTGAAGTCTGCTAGCTGACGAGGCTCCAAACCACTACT 5010
QY 841 TCAGAAATTCATCT--ATTATATAGATGAATTTATCTCTATATATATGCCCCGAAATTA 897

```

```

DB 5009 TCAGAAATTCATCTTAATATATAGATGAATTTATCTCTATATATGCCCCGAAATTA 4950
QY 898 ATTGCGGATCTGGAATTTATCATTTTATCAAAACATGATGATGATACAGAT 957
DB 4949 ATTGCGGATCTGGAATTTATCATTTTATCAAAACATGATGATGATACAGAT 4890
QY 958 GCGAACCATTGTGAATTTGTCATTCGAGTTATTAATCAAAATGTTAAATTAATCTTTA 1017
DB 4889 GCGAACCATTGTGAATTTGTCATTCGAGTTATTAATCAAAATGTTAAATTAATCTTTA 4830
QY 1018 CGAAGAGAGTGA 1029
DB 4829 CGAAGAGAGTGA 4818

RESULT 7
T84023
ID T84023 standard; DNA; 1269 BP.
XX
AC T84023;
XX
DT 01-SEP-1998 (first entry)
XX
DE DNA encoding a Riboflavin biosynthesis protein RIBG.
XX
KW Staphylococcus aureus proteini; ribozyme; antisense sequence; control;
KW Staphylococcal gene; regulatory element; bacterial gene expression;
KW vaccine; Staphylococcal infection; food poisoning; scalded skin syndrome;
KW toxic shock syndrome; ss.
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FT CDS 734..1033
FT /tag= a
XX
XX WO9730070-A1.
XX
XX 21-AUG-1997.
XX
XX 19-FEB-1997; 97WO-US02318.
XX
XX 20-FEB-1996; 96US-0011888.
XX
XX (SMIK ) SMITHKLINE BEECHAM CORP.
XX
XX Black MT, Burnham MK, Hodgson JE, Knowles DJC, Nicholas RO;
XX Pratt JM, Reichard RW, Rosenberg M, Ward JM;
XX WPI; 1997-424969/39.
XX P-PSDB; W28084.
XX
XX Novel polypeptide(s) from Staphylococcus aureus strain WCUH29 - used
XX to isolate antimicrobial compounds, and in vaccines against S.
XX aureus infection
XX
XX Claim 9; Page 825; 989pp; English.
XX
XX The present sequence encodes a Staphylococcus aureus protein, that,
XX based on homology with an Bacillus subtilis protein, is believed to be a
XX Riboflavin biosynthesis protein RIBG. The present sequence was obtained
XX from a library of clones of S. aureus WCUH 29 in Escherichia coli. The
XX DNA sequence can be used in the construction of ribozymes and antisense
XX sequences to control the expression of Staphylococcal genes. The DNA
XX sequence is also useful as a source of regulatory elements for the
XX control of bacterial gene expression. The encoded protein may be used
XX to produce vaccines to enable a host to produce specific antibodies
XX with antibacterial action. These vaccines and antibodies would protect
XX a host against invasion by S. aureus, and conditions relating to
XX Staphylococcal infection, e.g. Staphylococcal food poisoning, scalded
XX skin syndrome, and toxic shock syndrome.

```





```

DT 10-SEP-1999 (first entry)
XX
XX B. subtilis rib operon nucleotide sequence.
DE
XX Riboflavin: open reading frame; ORF: structural gene; promoter;
KW vitamin B2; Bacillus subtilis; rib operon; ds.
XX
OS Bacillus subtilis.
XX
XX Key Location/Qualifiers
FH 1..5567
FT CDS
FT
FT /*tag= a
FT /note= "reading frame 1 translated protein (Y21801)"
FT CDS
FT /*tag= b
FT /note= "reading frame 2 translated protein (Y21802)"
FT CDS
FT /*tag= c
FT /note= "reading frame 3 translated protein (Y21803)"
XX
XX US5925538-A.
XX
XX 20-JUL-1999.
XX
XX 24-AUG-1998; 98US-0138775.
XX
XX 11-SEP-1990; 90US-0581048.
XX 22-JUN-1989; 89US-0370378.
XX 21-APR-1992; 92US-0873572.
XX 06-FEB-1995; 95US-0384626.
XX 24-AUG-1998; 98US-0138775.
XX
XX (HOFF ) ROCHE VITAMINS INC.
XX
XX Erdenberger T, Hatch RT, Hermann T, Perkins JB;
PI Pero JG, Sloma A;
XX
XX WPI: 1999-418271/35.
DR P-PSDB; Y21801, Y21802, Y21803.
XX
XX Use of bacterial strains for the over production of riboflavin
PT
XX
XX Example 2; Fig 3A-S; 61pp; English.
XX
XX The invention relates to a method of producing riboflavin that comprises,
CC providing a recombinant bacterium containing an exogenously introduced
CC bacterial ORF (open reading frame) 3 (or ORFs) structural gene and an
CC exogenous promoter, culturing the bacterium, and recovery of the product.
CC The method is useful for the production of large quantities of riboflavin
CC (vitamin B2). The present sequence represents the complete nucleotide
CC sequence of B. subtilis rib operon (also see X81948 for the antisense
CC sequence).
XX
XX Sequence 5567 BP; 1751 A; 1096 C; 1297 G; 1423 T; 0 other;
SQ

```

```

Query Match 11.7%; Score 120.4; DB 20; Length 5567;
Best Local Similarity 51.0%; Pred. No. 7.6e-19;
Matches 317; Conservative 0; Mismatches 296; Indels 9; Gaps 1;

```

```

QY 1 ATGGATTTATGCGATTCACACTTGCATATATGATGACAAAGTCAACAGGTATTATCCACCC 60
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1116 atgaagctgacctagatcttgcaagcagcgcaagcagacacccaatccatccgctc 1175
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 61 GTTGGCGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1176 gtccgcgcctgtgtcgttaagcagcacaatctgcgaatg99gcccatttaaatat 1235
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 121 GGTGCAACGATCGGAGGTTCACACACTTGTATGTCGACAAACAATGCTGAAGTGGC 180
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1236 ggtgaagcctcagcagaatcgtcatcgtcatcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1295
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 181 ACGATTATATACGTTAGAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

DB 1296 gacattacgttacactcgaaccgctgcagccacttacggaataaacacccgcatgtcagaa 1355
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 241 AAAATTTATGATTTGATTAAGTATACCAAAAGTATACGCAACAAAGA-----CAAT 291
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1356 ttgattatcaacttcggtatcaaaagagctgcgtgcgagatgagagatcctaaccgctt 1415
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 292 TCGTTAGACACACATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 351
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1416 gtggtcgtgaaagaggtatcagcatgtatgaagaagctgcatgtgaggtaaagagcctc 1475
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 352 GATGATGAACGGGCAATCAATTTATACCAAGCTTTTAAAGCAACAAAGCAACGACTG 411
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1476 ctgcagcaccagcggaagagctgcatgtatgaataatcttcgcatcttatgagcagcgctt 1535
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 412 CCACAAATTTACAGTGAAGTATCTCAAGTGTATGATGATGATGATGATGATGATGATGATGAT 471
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1536 ccgtacgtcagcttaaaagcgctgcagcccttcgacgcaagaatgctcaccgacggtt 1595
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 472 CAAAGTCAATGATTTACTTAACAAAGAGTTAAACAGATGCTTATTAAGTAAAGATCGA 531
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1596 gacagcaaatgcatcagctcagagagctgcaagaacagatgctcagcaatatacagaaaca 1655
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 532 CACGACGACGTGTTACTGTAAGACGTACACTTGAATTGATGATGATGATGATGATGATGATGAT 591
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1656 caaccaagcatcttgcgtgagctgcaagctgcaagctgcaagctgcaagctgcaagctgca 1715
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 592 CGTATTCAAGATGAAAAAAC 613
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1716 agactgcggaatgtatacaaaaac 1737
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

RESULT 11
X81948/c
ID X81948 standard; DNA; 5567 BP.
XX
XX X81948;
XX
XX 10-SEP-1999 (first entry)
XX
XX B. subtilis rib operon complementary nucleotide sequence.
DE
XX
XX Riboflavin: open reading frame; ORF: structural gene; promoter;
KW vitamin B2; Bacillus subtilis; rib operon; complementary; ds.
XX
XX Bacillus subtilis.
XX
XX Key Location/Qualifiers
FH 1..5567
FT CDS
FT /*tag= a
FT /note= "reading frame 1 translated protein (Y21804)"
FT CDS
FT /*tag= b
FT /note= "reading frame 2 translated protein (Y21805)"
FT CDS
FT /*tag= c
FT /note= "reading frame 3 translated protein (Y21806)"
XX
XX US5925538-A.
XX
XX 20-JUL-1999.
XX
XX 24-AUG-1998; 98US-0138775.
XX
XX 11-SEP-1990; 90US-0581048.
XX 22-JUN-1989; 89US-0370378.
XX 21-APR-1992; 92US-0873572.
XX 06-FEB-1995; 95US-0384626.
XX 24-AUG-1998; 98US-0138775.
XX
XX (HOFF ) ROCHE VITAMINS INC.
XX
XX Erdenberger T, Hatch RT, Hermann T, Perkins JB;
PI

```

PI Pero JG, Sioma A;  
 XX WPI: 1999-418271/35.  
 DR P-PSDB: Y21804, Y21805, Y21806.  
 XX  
 XX  
 PT Use of bacterial strains for the over production of riboflavin  
 XX  
 PS Example 2; Fig 3A-S; 61pp; English.  
 XX  
 CC The invention relates to a method of producing riboflavin that comprises,  
 CC providing a recombinant bacterium containing an exogenously introduced  
 CC bacterial ORF (open reading frame) 3 (or ORF5) structural gene and an  
 CC exogenous promoter, culturing the bacterium, and recovery of the product.  
 CC The method is useful for the production of large quantities of riboflavin  
 CC (vitamin B2). The present sequence represents the complementary  
 CC nucleotide sequence of B. subtilis rib operon.  
 XX  
 SO Sequence 5567 BP; 1422 A; 1296 C; 1097 G; 1752 T; 0 other;

Query Match 11.7%; Score 120.4; DB 20; Length 5567;  
 Best Local Similarity 51.0%; Pred. No. 7.6e-19;  
 Matches 317; Conservative 0; Mismatches 296; Indels 9; Gaps 1;

QY 1 ATGCATTATCGCATTCACCTTGCATAATATGCTACACGCTCAACAGCTGTTATCCACC 60  
 DB 4452 ATGAAGCTGGCCTTAGTCTTGCAGACGAGGCGGACAGACCGAATCCCATCCGCTC 4393  
 QY 61 GTGGCGCCTGTGTAGTTATGAAGTAGGATTGTTGTTGTTGTTGTTGTTGTTGTTGTT 120  
 DB 4392 GTGGCGCCTGT 4333  
 QY 121 GGTGCAAGCATGGGAGGTTCAAGCACTGTATATGTCACACAAATGCTGAAGTGGC 180  
 DB 4332 GGTGAACCTCATGTGAGAAATTCATCCATATGCTGTGAGACATGACAGAGGTGCC 4273  
 QY 181 ACATTTATTTAGTTAGTAGCCCATGTAGTATTTTGTTCACACCACTGTTTAAAC 240  
 DB 4272 GACATTTATGTTACCTGACCTGACGATTCAGCAAAACACCGCATGTGCAGAA 4213  
 QY 241 AAAATTTATGTTATGTAATGACAAAGTATATGCAACAAAGA-----CAAT 291  
 DB 4212 TTGATTTATCAACTCTGTGTATCAAAAGAGTGTCTGTGCGATGACAGATCCATCCGCTT 4153  
 QY 292 TCGTTAGACACACATGCTGTATGACGCTTACGGGCTCACGGTATTTAGTTGCGTT 351  
 DB 4152 GTGGCTGGAAGAGGATCAGCATGATGAAGAAGCTGCAATTGAGTGAAGGAAGCATC 4093  
 QY 352 GATGATGACAGGCGCATCACAATTTATACCAAGACTTTTAAAGCAAAAGCAAACTG 411  
 DB 4092 CTGGCAGACCGGCGGAGAGGCTGAATAAATTTCTGCACCTTTATGAGGACAGGCTT 4033  
 QY 412 CCACAATTTACAGTGAAGATATCTGCAAGTTTATGATGTTAAACAGGATGATATGA 471  
 DB 4032 CCGTAGCTCACGCTAAAGGCGCTGCGCTTGACGCAAGATAGCTACACGACGCGT 3973  
 QY 472 CAATGATGATGATTTACTTAACAAGAGTTAAACAGAGTCTATTAAGTTAAGACATGA 531  
 DB 3972 GACAGCAATGATGATCAGTCAAGGCTGCAAGAGCAGATCTCAGCAATACAGAAACA 3913  
 QY 532 CACGACGCACTGTTAACTGACAGACGTCAGTGAATTTAGATGCCAATATACTACA 591  
 DB 3912 CACCAAGCATTTTATGTCGACAGTGGCACAGTGAAGCCGACATCCGAGCTTAACCTGC 3853  
 QY 592 CGTATTCAGATGAGAAAAACC 613  
 DB 3852 AGACTGCCGATGTACAAAAAC 3831

RESULT 12  
 V71600  
 ID V71600 standard; DNA: 5567 BP.  
 XX

AC V71600;  
 XX  
 DT 01-MAR-1999 (first entry)  
 XX  
 DE Bacillus subtilis rib operon.  
 XX  
 KW Riboflavin; vitamin B2; riboflavin synthase; deaminase; reductase;  
 XX GTP cyclohydrolase; rib operon; ds.  
 XX  
 OS Bacillus subtilis strain 168.  
 XX  
 FH location/Qualifiers  
 FT key 352..357  
 FT RBS /tag= a  
 FT CDS 364..681  
 FT /tag= b  
 FT /label= ORF6  
 FT 708..749  
 FT terminator /tag= c  
 FT /note= "rho-independent transcription terminator"  
 FT promoter 771..799  
 FT /tag= d  
 FT /label= P1  
 FT 1034..67  
 FT /tag= e  
 FT /note= "rho-independent transcription terminator"  
 FT RBS 1184..1195  
 FT /tag= f  
 FT 1422..1433  
 FT /tag= g  
 FT CDS 1437..2186  
 FT /tag= h  
 FT /label= ORF5  
 FT 2178..2194  
 FT /tag= i  
 FT 2197..2844  
 FT /tag= j  
 FT /label= ORF4  
 FT 2528..2556  
 FT /tag= k  
 FT /label= P2  
 FT 2843..2854  
 FT /tag= l  
 FT CDS 2859..4055  
 FT /tag= m  
 FT /label= ORF3  
 FT 4071..4085  
 FT /tag= n  
 FT CDS 4088..4552  
 FT /tag= o  
 FT /product= riboflavin synthase beta subunit  
 FT 4545..4574  
 FT /tag= p  
 FT /label= P3  
 FT 4653..4659  
 FT /tag= q  
 FT CDS 4665..5039  
 FT /tag= r  
 FT /label= ORF2  
 FT 5038..5090  
 FT /tag= s  
 FT /note= "rho-independent transcription terminator"  
 FT complement (5053..5567)  
 FT /tag= t  
 FT /label= ORF1  
 XX  
 XX US5837528-A.  
 XX  
 PD 17-NOV-1998.  
 XX  
 PF 06-FEB-1995; 95US-0384626.  
 XX  
 PR 11-SEP-1990; 90US-0581048.

BR 22-JUN-1989; 89US-0370378.  
 PR 21-APR-1992; 92US-0873572.  
 PR 06-FEB-1995; 95US-0384626.  
 XX  
 XX (HOFF ) HOFFMANN LA ROCHE INC.  
 PI Perkins JB, Pero JC, Sloma A;  
 DR WPI: 1999-023456/02.  
 DR P-PSDB; W81549-55.  
 XX  
 PT Nucleic acid encoding proteins involved in bacterial synthesis of  
 PT riboflavin - also mutant or recombinant bacteria producing  
 PT riboflavin at high levels  
 XX  
 PS Example 7; Fig 3A-S; 62pp; English.  
 XX

CC This is the nucleotide sequence of the rib operon of *Bacillus*  
 CC subtilis. The rib operon includes an open reading frame (ORF)  
 CC encoding the beta-subunit (see W81553) of riboflavin synthase, and  
 CC 6 other ORFs, of which ORF1 is a partial coding region encoding a  
 CC 170-amino acid polypeptide (see W81555), ORF2 encodes a putative  
 CC rib-specific reductase (see W81554), ORF3 encodes a putative GTP  
 CC cyclohydrolase (see W81552), ORF4 encodes the riboflavin synthase  
 CC alpha subunit (see W81551), ORF5 encodes a putative rib-specific  
 CC deaminase (see W81550) and ORF6 encodes a 38 kDa protein (see  
 CC W81549) of unknown function. The invention provides vectors and  
 CC recombinant bacteria for overproducing riboflavin, in which nucleic  
 CC acids overproducing riboflavin biosynthetic proteins are introduced  
 CC into the chromosome of the host organism at multiple sites or in  
 CC multiple copies. The rib operon, e.g. comprising ORFs 2-5 and  
 CC the beta-riboflavin synthase gene, is used to make such recombinant  
 CC bacteria. The rib operon can be deregulated by replacing its  
 CC control regions with sequences that allow constitutive or  
 CC unregulated expression. The recombinant bacteria produce at least  
 CC 10 g/l riboflavin.  
 SO

Sequence 5567 BP; 1749 A; 1102 C; 1298 G; 1418 T; 0 other;

Query Match 11.7%; Score 120.4; DB 20; Length 5567;  
 Best Local Similarity 51.0%; Pred. No. 7.6e-19;  
 Matches 317; Conservative 0; Mismatches 296; Indels 9; Gaps 1;

QY 1 ATGATATATGCGATTCACCTGCAATATATGTACAGGTCAACAGGTGTAATCCACC 60  
 DB 1116 atgaagctgagccttagatcttggaagcggagcagcagcagcagcagcagcagc 1175  
 QY 61 GTTGGCGCTGTGTATGATTAAGTAGAGATTGTTGATGTCGACACTTGAGAAA 120  
 DB 1176 gtgcgctgctgt 1235  
 QY 121 GGTGACAGCATGCGAGGTTGACGACTTGTATATGCGACCAAAATGCTGAGGTGG 180  
 DB 1236 ggtgaagcctcagcagaagctatcgcacatagctgagcagcagcagcagcagcagc 1295  
 QY 181 ACGATTATATATGCTTATAGAGCATGATGATTTGTTGTTCAACACCACTGTGTTAAC 240  
 DB 1296 gacattcagctacactcgaacgcgcagcagcagcagcagcagcagcagcagcagc 1355  
 QY 241 AAAATTATGATGTAAGTATACCAAAAGTATACGCAACAAAGA-----CAAT 291  
 DB 1356 ttgattatcaactctggtatcaaaagagtgctgctgagcagatgagatcttaaccgctt 1415  
 QY 292 TCGTTAGACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 351  
 DB 1416 gtgcgctggaagggagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1475  
 QY 352 GATGATGAACGGGATCACAATATATACCAAGACTTTTAAAGCAAAAGCAAGCAACTG 411  
 DB 1476 ctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1535  
 QY 412 CCACAATATACAGTGAAGTATCTGCAAGTTAGTGTAAACCAAGCAATGATATGGA 471

DB 1536 ccgtacgtcagcctaaagcggctgcgcagccttgacggcagcagcagcagcagcagc 1595  
 QY 472 CAAAGTCATGATGATTAACCAAGAGTTAAACAAAGATGCTGTATAGTTAGACATGCA 531  
 DB 1596 gacacgaatgcatcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1655  
 QY 532 CACGACGAGTGTAACTGGAAGAGCTACAGTTGATTAAGATGATCCACAATATCTACA 591  
 DB 1656 caccgaagcatttagctgagctgagcagcagcagcagcagcagcagcagcagcagcagc 1715  
 QY 592 CGTATTCAGATGGAATAAAC 613  
 DB 1716 agactgcgcaatgtaacaaac 1737

RESULT 13  
 293752  
 ID 293752 standard; DNM: 5567 BP.  
 XX  
 AC 293752;  
 XX  
 DT 16-AUG-2000 (first entry)  
 XX  
 DE Rib operon of *Bacillus subtilis*.  
 XX  
 KW Riboflavin; rib operon; yeast; bacteria; transformation;  
 KW transcription element; gene expression; hair loss; skin inflammation;  
 KW growth disorder; vision; vision; sight; ds.  
 XX  
 OS *Bacillus subtilis*.  
 XX  
 FH Key  
 FT CDS  
 FT Location/Qualifiers  
 FT 1..5567  
 FT /tag= a  
 FT /note= "Contains in frame stop codons"  
 FT  
 FT RBS  
 FT 62..68  
 FT /tag= b  
 FT 363..680  
 FT /tag= c  
 FT /label= ORF 6  
 FT 708..748  
 FT /tag= d  
 FT /label= Rho independent transcription terminator  
 FT 771..799  
 FT /tag= e  
 FT /label= p1 promoter  
 FT 1034..1067  
 FT /tag= f  
 FT /label= Rho independent transcription terminator  
 FT 1084..1095  
 FT /tag= g  
 FT 1101..2186  
 FT /tag= h  
 FT /label= ORF5  
 FT 1422..1433  
 FT /tag= i  
 FT 2177..2194  
 FT /tag= j  
 FT 2197..2844  
 FT /tag= k  
 FT /label= ORP4  
 FT 2528..2556  
 FT /tag= l  
 FT /label= p2 promoter  
 FT 2843..2854  
 FT /tag= m  
 FT 2859..4055  
 FT /tag= n  
 FT /label= ORF3  
 FT 4071..4085  
 FT /tag= o  
 FT 4088..4552  
 FT CDS

```

FT      /tag- p
FT      /product= Beta riboflavin synthase
FT      4545..4574
FT      promoter
FT      /tag- g
FT      /label= p3 promoter
FT      4653..4659
FT      RBS
FT      /tag- t
FT      4665..5039
FT      CDS
FT      /tag- s
FT      /label= ORF2
FT      5038..5090
FT      terminator
FT      /tag- t
FT      /label= Rho independent transcription terminator
FT      complement (5053..5567)
FT      /tag- u
FT      /label= ORF1
FT      CDS
FT      EP1001026-A2.
FT      17-MAY-2000.
FT      22-JUN-1990; 99EP-0124888.
FT      22-JUN-1989; 89US-0370378.
FT      22-JUN-1990; 90EP-0111916.
FT      (HOFF ) HOFFMANN LA ROCHE & CO AG F.
FT      Perkins JB, Pero JG, Sloma A:
FT      WPI; 2000-320439/28.
FT      P-PSDB; Y83269, Y83270, Y83271.
FT      Novel vector, used for the large scale production of riboflavin,
FT      comprises a bacterial or yeast riboflavin biosynthetic protein under
FT      the control of a foreign transcription element
FT      Claim 1; Figure 3; 66pp; English.
FT      Vectors comprising a nucleic acid sequence of bacterial or
FT      yeast origin, coding for one or more riboflavin biosynthetic proteins,
FT      and one or more transcription elements not naturally associated with
FT      the nucleic acid sequence are new. The vector and bacteria transformed
FT      with it are used for the large scale production (over 10g/l) of
FT      riboflavin. The riboflavin produced can be used to treat disorders
FT      associated with riboflavin deficiency such as hair loss, skin
FT      inflammation, vision deterioration and growth failure.
FT      Sequence 5567 BP; 1751 A; 1097 C; 1296 G; 1423 T; 0 other;
SQ

```

Query Match 11.7%; Score 120.4; DB 21; Length 5567;  
 Best Local Similarity 51.0%; Pred. No. 7.6e-19;  
 Matches 317; Conservative 0; Mismatches 296; Indels 9; Gaps 1;

```

QY      1 ATGCATTATGCGATTCACTTGCATAATATGTCACAAGTCACAAGCGTGTTAATCCACCC 60
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      1116 atgaagctggccttagatctcggaagcagggcgaagacagaccgcatccatccgctc 1175
QY      61 GTTGGCGCTTTGTAGTTAATGAAGTAGATTGTTGGTTGTCACACTTGAGAAAA 120
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      1176 gtcgagcgtctgtctgaagacggaacattgtcgaatgagcgccatttaaatat 1235
QY      121 GGTGACAGACGTCGGAGGTTCAGCACTTGATATGTCACACAAATATGCTGAAGTGGC 180
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      1236 ggtgaagctcatgctgagagctcatgcatcatatgctggaacatgcaagaggtgccc 1295
QY      181 ACATTATATTAATGCTAGAGCAGCATGTAGTATTTGTTCAACACCCACCCCTGTGTAA 240
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      1296 gacattacgttcaacacgcgacgtgacgcatcgaagaaaaaacacgcgcatgtgcaga 1355
QY      241 AAATTATTCATTTGTAATGAGCAAAAGTGTATGACCAACAAAAGA-----CAAT 291
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

DB      1356 ttgatctcaactctggtatcaaaaagtgcttctggtgcgatgagagatcctaaccgctt 1415
QY      292 TCGTTAGACACACATGCTGATGAGACCGCTTACGGGCTCACGGTATTTAGATTGAATGCGTT 351
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      1416 gtgctctggaagagagatcagcatgatgaagaagctgtgcatctgagtggaaggaagcgtc 1475
QY      352 GATGATGACCGGCGATCACAAATTATVACCAAGACTTTTTTAAAGCAAAAGCAACCAACTG 411
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      1476 ctgcagaccagcggagagagctgatcatgaaaaattcttcgactttatgaagacagcctt 1535
QY      412 CCACAAATTACAGTGAAGATATCTGCAAGTTTACATGTTAAACAGCAATGATTAATGA 471
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      1536 ccgtacgtccacgcgcaaaagcgcctgcagcccttcgcgcgaagatagctaccagcagcgtt 1595
QY      472 CAAAGTCATGATGATTACTTAACAAGAGTTAAACAAATGCTCTATTAAGTAAACATCGA 531
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      1596 gacagcaaatggtatcagtcagcagagcgcgcaagaagatgctcagcaataacgaaacaa 1655
QY      532 CACGACGCGAGTGTAACTGGAAGACGTACAGTTGAATTAGATCCACATATATCTACA 591
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      1656 caccaaagcatltagtcgagtggtgcacagtggaagccgacaaatccgagcttaacctgc 1715
QY      592 CGTATTCAGATGCAAAAAC 613
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      1716 agactgccgaatgtaacaaaac 1737

RESULT 14
V52173/C
ID      V52173 standard; DNA; 14273 BP.
XX
XX      V52173;
XX
DE      23-OCT-1998 (first entry)
XX
XX      Streptococcus pneumoniae genome fragment SEQ ID NO:40.
XX
XX      Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;
XX      computer readable medium; vaccine; pharmaceutical composition; ds.
XX
XX      Streptococcus pneumoniae.
XX
XX      W09818931-A2.
XX
XX      07-MAY-1998.
XX
XX      30-OCT-1997; 97WO-US19588.
XX
XX      31-OCT-1996; 96US-0029960.
XX
XX      (HUMA-) HUMAN GENOME SCI INC.
XX
XX      Barash SC, Choi GH, Dillon PJ, Dougherty BA, Fannon M:
XX      Kunsch CA, Rosen CA;
XX      WPI; 1998-272225/24.
XX
XX      Computer-readable medium with recorded Streptococcus pneumoniae
XX      polynucleotide sequences - useful in diagnostic kits and assays, and
XX      pharmaceutical compositions and vaccines for Streptococcus
XX      pneumoniae
XX
XX      Claim 1; Page 389-398; 1409pp; English.
XX
XX      The present invention describes a computer readable medium which has
XX      the nucleotide sequences SEQ ID NO:1 to 391 (V52134 to V52524) recorded
XX      on it, or a representative fragment or a sequence at least 95% identical
XX      to SEQ ID NO: 1 to 391. The nucleotide sequences depicted in SEQ ID NO:1
XX      to 391 (V52134 to V52524) are genomic fragments from Streptococcus
XX      pneumoniae. The present invention also describes an isolated nucleic acid
XX      molecule encoding a homologue of any of the fragments of the S.pneumoniae
XX      genome (SEQ ID NO:1 to 391) where the nucleic acid molecule is produced
XX      by a process comprising: (a) screening a genomic DNA library using as a

```

CC probe a target sequence defined by any of the sequences in SEQ ID NO:1  
 CC to 391, identifying members of the library which contain sequences  
 CC that hybridize to the target sequence and isolating the nucleic acid  
 CC molecules from the members; or (b) isolating mRNA, DNA or cDNA produced  
 CC from an organism, amplifying nucleic acid molecules whose nucleotide  
 CC sequence is homologous to amplification primers derived from the  
 CC fragment of the S. pneumoniae genome to prime the amplification and  
 CC isolating the amplified sequences. The computer readable medium can be  
 CC used in a computer-based system for identifying fragments of the  
 CC S. pneumoniae genome of commercial importance, or expression modulating  
 CC fragments of the S. pneumoniae genome. Products from the present  
 CC invention can be used in diagnosis kits and assays, and pharmaceutical  
 CC compositions and vaccines for S. pneumoniae.

XX Sequence 14273 BP; 4357 A; 2428 C; 2694 G; 4794 T; 0 other;

Query Match 11.1%; Score 114.2; DB 19; Length 14273;  
 Best Local Similarity 49.3%; Pred. No. 2.7e-17;  
 Matches 387; Conservative 0; Mismatches 388; Indels 10; Gaps 3;

OY 1 ATGATTTATGCGATTCACCTTGCATATATGATACAGGTCAACAGGTGTATCCACC 60  
 DB 10928 ATGAATTTAGCATATAAACTGCGACAAAAGGGCGTGTACGTCAATCCCATCTATG 10869  
 OY 61 GTTGGCGCTGTGTATGATGAGGTAGATTTGTTGTTATGTCACACTTGTAGAGAAA 120  
 DB 10868 GTCCGCGCATTTATGTAAGATATATCATTTATCGCAGGTATCATGATGTTTTT 10809  
 OY 121 GGTGACAGCATGGGAGGTTCAACGCTGATGCGACACAAATGCTGAAGTGGC 180  
 DB 10808 GTGGCCCATCATGTGAGAGAAATGCTCTTAAACCTGAGAAATCCCTTCGAGCG 10749  
 OY 181 ACGATTATATTTACGTTAGAGCCATGTAGTCATTTGTTGTTCAACCACTGTGTAAAC 240  
 DB 10748 ACGTTTATGTACACTGGAACCGTGTCTCTCGGSAACACCTCCGTATAGAT 10689  
 OY 241 AAATTTATGATTTGATGATGACCAAAAGTATGATACGCAACAAAGA-----CAATT 292  
 DB 10688 GCTATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 10629  
 OY 293 CGTTAGACACACATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 352  
 DB 10628 GTATCTGAAAAAGAGTAAAGTACTTGAGCAAAATATCTTCAAGTTACTTGGAAAT 10569  
 OY 353 ATGATGAACGGGCATCACATATTTACCAAGACTTTTAAAGCA-AAAGCAAGCAACTG 411  
 DB 10568 TTAGAAATTTGAGTGTCTTAATTAATAAAAGTTTAAAGTATATTAATTAATTAAT 10509  
 OY 412 CCACAATTTACAGTGAAGTATCTCAAGTTTATGATGTAACCAAGCCAGTATATGGA 471  
 DB 10508 CCGTATGTTTTTATGAAATATGCAATGTCATGTAAGTGAAGTGAAGTGAAGTGA 10449  
 OY 472 CAAAGTCAATGATTTACTTAAACAAAGAGTTAAACAAGTCTTAAAGTTAAGTATGCA 531  
 DB 10448 CAATCAATGATTTACTTGAAGAGAGCAAGCAAGTGTGATGATGATGATGATGATG 10389  
 OY 532 CACGACGAGTGTAACTGAGAGAGTACAGTGTGAATTAGATGATGATGATGATGATG 591  
 DB 10388 GTTACTGCAATTTATGTTGAGATCATATCTTTATTAAGAGAGATCTTTTGTGACAT 10329  
 OY 592 CGTATTTCAAGATGAAAAAACCCTTAAAGTATTTGTCTTAACTGCTGGGAATATTCAT 651  
 DB 10328 AGATTGGAGGAGGAAAAAACCCTTAAAGTATTTGTGATGATGATGATGATGATG 10269  
 OY 652 TTTA-ATGAGCAATTTTAAAGATGATGATGATGATGATGATGATGATGATGATG 710  
 DB 10268 CTTTACTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 10209  
 OY 711 AAATTTAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 770  
 DB 10208 GAAGCAAAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 10149

OY 771 AACAA 775  
 DB 10148 AAAAA 10144

RESULT 15

X77179 standard; DNA; 1101 BP.

X77179;

04-AUG-1999 (first entry)

S. pneumoniae ribG polypeptide encoding DNA.

Streptococcus pneumoniae; pyrimidine deaminase; pyrimidine reductase;

ribG; immune response; antimicrobial; H. pylori infection; cancer;

gastrointestinal carcinoma; gastric ulcer; gastritis; ss.

Streptococcus pneumoniae.

WO927126-A1.

03-JUN-1999.

23-NOV-1998; 98WO-US25010.

25-NOV-1997; 97US-0979616.

(SMK ) SMITHKLINE BEECHAM CORP.

Black MT, Debonou C, Fedon JC, Hodgson JE, Jaworski JD;

Knowles DJC, Kosmatka AL, Lonetto MA, Mooney J, Nicholas RO;

Palmer LM, Shilling LK, Stodola RK, Wang M, Warren RL;

Zhong Y;

WPI: 1999-347726/29.

P-PSDB: Y117949.

New Streptococcus pneumoniae ribG polypeptide and polynucleotide

useful in the treatment of gastric ulcer and gastritis

Claim 6; Page 5; 48pp; English.

The invention relates to Streptococcus pneumoniae pyrimidine deaminase

and pyrimidine reductase (ribG). S. pneumoniae ribG and its antagonists a

re used to treat individuals in need of the polypeptide. Disease related

to expression or activity of ribG can be determined by analysing the

nucleic acid sequence encoding ribG or detecting the ribG polypeptide in

a sample. ribG can also be used to identify antagonists or agonists.

CC ribG, or its related nucleic acid, also has use as a vaccine to induce

an immunological response in an animal. Antimicrobial compounds (e.g.

agonists and antagonists of ribG), especially broad-spectrum antibiotics,

may be of use in the treatment of Helicobacter pylori infection. This

CC may decrease the advent of H. pylori-induced cancers, such as

CC gastrointestinal carcinoma. The treatment should also cure gastric ulcers

CC and gastritis. The present sequence represents a DNA encoding the

CC S. pneumoniae ribG polypeptide.

Sequence 1101 BP; 392 A; 180 C; 216 G; 313 T; 0 other;

Query Match 10.9%; Score 112.6; DB 20; Length 1101;  
 Best Local Similarity 49.2%; Pred. No. 3.2e-17;  
 Matches 386; Conservative 0; Mismatches 389; Indels 10; Gaps 3;

OY 1 ATGATTTATGCGATTCACCTTGCATATATGATACAGGTGTATCCACC 60  
 DB 19 atgaatgacataaaactgcacaaaagggctgttgaatgcacacatccatctatg 78  
 OY 61 GTTGGCGCTGTGTATGATGAGGTAGATTTGTTGTTATGTCACACTTGTAGAAA 120  
 DB 79 gctggcgaattatcgttaaaagataatcaatcgcgaaggtatcatcaggttttc 138



```
OY 121 GGTGACAGCATCGGAGTTCAAGCATTGATATGGCACACAAAATGCTGAAGTGGC 180
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 139 ggtggccacatctgagagaaatgctcttaaaaactgtagaaatccctgctggagcg 198
OY 181 ACGATTTATTTAGCTTAGAGCCATGATGATTTTGGTTCAACACCACCTGTGTAAAC 240
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 199 agccttaattgaaccttgtaacctgtgtcacttcggaaacacacctcctgtatagat 258
OY 241 AAAATTATTGATTGTAAGATAGCAAAAGTATACGACACAAAAGA-----CAATT 292
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 259 gctataatcgatagtgatatacaagagtagtcaatggaagccttagacttaatccatt 318
OY 293 CGTTAGACACACATGCTGATGAGAGCTTACGGGCTCACGGTATTGAGGTTGAATGCTTG 352
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 319 gtatctgaaaagagagaaagatcttgagaaataatctcaagtlactgttggatt 378
OY 353 ATGATGACCGGCATCATATTATACCAAGACTTTTAAAGCA-AAAGCAAGCAACTG 411
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 379 ttgaaaaatgagtgcttaacttaataaaaagtttagaaagtatactaccagcatgta 438
OY 412 CCACAAATTACAGTGAAGATCTGCAAGTGTAGATGTTAAACAGCAATGATATGA 471
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 439 ccctatgttcttgaataatgtcaatgtaatggaataaataagccactaaacaat 498
OY 472 CAAAGTCAATGATTAAGTAAAGAGGTTAAACAGATGCTATAAGTTAAGCATGA 531
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 499 caatccaatlgatctactgaagaagaagaagaagcatgtgcatcagttacgacactat 558
OY 532 CACGACGCACTGTTAACTGCAAGACGTACGTTGAATTAGATGATCCACAATTAATA 591
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 559 gtagtgcacatlaagtgaggagcaatcactgtaacagaagatcccttgcgacatg 618
OY 592 CGATTTCAGATGGAAGAAACCTATTAAGTAAAGTAAATTTGCTAAGTCTGGGAATATTCAT 651
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 619 agatctggagaaagaaataatccatccgatacatatgcatatacaatlaagaaacct 678
OY 652 TTTA-ATCAGCAAAATTATCAAGATGAATCAACACCAATTGGATATATCTGAAAATCC 710
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 679 cttaaccttaaatcgtaaaaacagcaatlgatatataaacttaacttgccactcct 738
OY 711 AAATTTAAGCAAGCATCAACATATTTGAATTATTTACTTGAAGTCTGTGATTTAAC 770
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 739 gaagacaaaaataaatgaagctatatcaaatcatatgctgtaaatlacttccataag 798
OY 771 AACAA 775
    ||| |
Db 799 aaaaa 803
```

Search completed: May 7, 2001, 12:43:42  
Job time: 4383 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using 'sw model

Run on: May 7, 2001, 10:20:14 ; Search time 1037.25 Seconds  
(without alignments)  
8666.614 Million cell updates/sec

Title: US-09-376-633-1

Perfect score: 1029  
Sequence: 1 ATGGATTATGCGATTCAACT.....TAACCTTAGAAGAAGTGA 1029

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 9623517 seqs, 4368049070 residues

Total number of hits satisfying chosen parameters: 19247034

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*  
10: gb\_est10:\*  
11: gb\_est11:\*  
12: gb\_est12:\*  
13: gb\_est13:\*  
14: gb\_est14:\*  
15: gb\_est15:\*  
16: gb\_est16:\*  
17: gb\_est17:\*  
18: gb\_est18:\*  
19: gb\_est19:\*  
20: gb\_est20:\*  
21: gb\_est21:\*  
22: gb\_est22:\*  
23: gb\_est23:\*  
24: gb\_est24:\*  
25: gb\_est25:\*  
26: gb\_est26:\*  
27: gb\_est27:\*  
28: gb\_est28:\*  
29: gb\_est29:\*  
30: gb\_est30:\*  
31: gb\_est31:\*  
32: gb\_est32:\*  
33: gb\_est33:\*  
34: gb\_est34:\*  
35: gb\_est35:\*  
36: gb\_est36:\*  
37: gb\_est37:\*  
38: gb\_est38:\*  
39: gb\_est39:\*  
40: gb\_est40:\*  
41: gb\_est41:\*  
42: gb\_est42:\*  
43: gb\_est43:\*

44: em\_esthum10:\*  
45: em\_esthum11:\*  
46: em\_esthum12:\*  
47: em\_esthum13:\*  
48: em\_esthum14:\*  
49: em\_esthum15:\*  
50: em\_esthum16:\*  
51: em\_esthum17:\*  
52: em\_esthum18:\*  
53: em\_esthum19:\*  
54: em\_esthum20:\*  
55: em\_esthum21:\*  
56: em\_esthum22:\*  
57: em\_esthum23:\*  
58: em\_esthum24:\*  
59: em\_esthum25:\*  
60: em\_esthum26:\*  
61: em\_esthum27:\*  
62: em\_esthum28:\*  
63: em\_esthum29:\*  
64: em\_esthum30:\*  
65: em\_esthum31:\*  
66: em\_esthum32:\*  
67: em\_esthum33:\*  
68: em\_esthum34:\*  
69: em\_esthum35:\*  
70: em\_esthum36:\*  
71: em\_esthum37:\*  
72: em\_esthum38:\*  
73: em\_esthum39:\*  
74: em\_esthum40:\*  
75: em\_esthum41:\*  
76: em\_esthum42:\*  
77: em\_esthum43:\*  
78: em\_esthum44:\*  
79: em\_esthum45:\*  
80: em\_esthum46:\*  
81: em\_esthum47:\*  
82: em\_esthum48:\*  
83: em\_esthum49:\*  
84: em\_esthum50:\*  
85: em\_esthum51:\*  
86: em\_esthum52:\*  
87: em\_esthum53:\*  
88: em\_esthum54:\*  
89: em\_esthum55:\*  
90: em\_esthum56:\*  
91: em\_esthum57:\*  
92: em\_esthum58:\*  
93: em\_esthum59:\*  
94: em\_esthum60:\*  
95: em\_esthum61:\*  
96: em\_esthum62:\*  
97: em\_esthum63:\*  
98: em\_esthum64:\*  
99: em\_esthum65:\*  
100: em\_esthum66:\*  
101: em\_esthum67:\*  
102: em\_esthum68:\*  
103: em\_esthum69:\*  
104: em\_esthum70:\*  
105: em\_esthum71:\*  
106: em\_esthum72:\*  
107: em\_esthum73:\*  
108: em\_esthum74:\*  
109: em\_esthum75:\*  
110: em\_esthum76:\*  
111: em\_esthum77:\*  
112: em\_esthum78:\*  
113: em\_esthum79:\*  
114: em\_esthum80:\*  
115: em\_esthum81:\*  
116: em\_esthum82:\*

117: gb\_est48:\*  
 118: gb\_est49:\*  
 119: gb\_est50:\*  
 120: gb\_est51:\*  
 121: gb\_est52:\*  
 122: gb\_est53:\*  
 123: gb\_est54:\*  
 124: gb\_est55:\*  
 125: gb\_est56:\*  
 126: gb\_est57:\*  
 127: gb\_est58:\*  
 128: gb\_est59:\*  
 129: gb\_est60:\*  
 130: gb\_est61:\*  
 131: gb\_est62:\*  
 132: gb\_est63:\*  
 133: gb\_est64:\*  
 134: gb\_est65:\*  
 135: gb\_est66:\*  
 136: gb\_est67:\*  
 137: gb\_est68:\*  
 138: gb\_est69:\*  
 139: gb\_est70:\*  
 140: gb\_est71:\*  
 141: gb\_est72:\*  
 142: gb\_est73:\*  
 143: gb\_est74:\*  
 144: gb\_est75:\*  
 145: gb\_est76:\*  
 146: gb\_est77:\*  
 147: gb\_est78:\*  
 148: gb\_est79:\*  
 149: gb\_est80:\*  
 150: gb\_est81:\*  
 151: gb\_est82:\*  
 152: gb\_est83:\*  
 153: gb\_est84:\*  
 154: gb\_est85:\*  
 155: gb\_est86:\*  
 156: gb\_est87:\*  
 157: gb\_est88:\*  
 158: gb\_est89:\*  
 159: gb\_est90:\*  
 160: gb\_est91:\*  
 161: gb\_est92:\*  
 162: gb\_est93:\*  
 163: gb\_est94:\*  
 164: gb\_est95:\*  
 165: gb\_est96:\*  
 166: gb\_est97:\*  
 167: gb\_est98:\*  
 168: gb\_est99:\*  
 169: gb\_est100:\*  
 170: gb\_est101:\*  
 171: gb\_est102:\*  
 172: gb\_est103:\*  
 173: gb\_est104:\*  
 174: gb\_est105:\*  
 175: gb\_est106:\*  
 176: gb\_est107:\*  
 177: gb\_est108:\*  
 178: gb\_est109:\*  
 179: gb\_est110:\*  
 180: gb\_est111:\*  
 181: gb\_est112:\*  
 182: gb\_est113:\*  
 183: gb\_est114:\*  
 184: gb\_est115:\*  
 185: gb\_est116:\*  
 186: gb\_est117:\*  
 187: gb\_est118:\*  
 188: gb\_est119:\*  
 189: gb\_est120:\*

Result	Score	Match	Length	ID	Description
1	70.4	6.8	1094	229	AL101513 Drosophila
2	68.2	6.6	420	229	AV426745 AV426745
3	63.2	6.1	1101	229	AL063921 Drosophila
4	63	6.1	469	229	BE920828 EST424597
5	62.4	6.1	1101	229	AL069706 Drosophila
6	61.8	6.0	1001	229	AL105023 Drosophila
7	61.8	6.0	1101	229	AL069706 Drosophila
8	61.2	5.9	1101	229	AL108721 Drosophila
9	60.2	5.9	1101	229	AL068307 Drosophila
10	59.6	5.8	1101	229	AL068307 Drosophila
11	59.2	5.8	554	137	BE600990 p11_95_F0
12	58.4	5.7	842	229	AL055007 Drosophila
13	58	5.6	781	229	AL053444 Drosophila
14	58	5.6	996	229	AL071063 Drosophila
15	56.8	5.5	1218	166	BE427889 PFR6899 I
16	56.6	5.5	1201	229	AL106652 Drosophila
17	56.2	5.5	1101	229	AL063932 Drosophila
18	56	5.4	945	231	AL285149 Tetradon

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

```
19 56 5 4 959 229 CNS00655 AL062806 Drosophila
20 55 5.3 532 201 A0051567 A0051567 dbx0002c
21 55 5.3 639 229 CNS0170D AL108367 Drosophila
22 54.8 5.3 337 214 A091504 A091504 Rf02461
23 54.4 5.3 1101 229 CNS0161V AL106801 Drosophila
24 53.8 5.2 522 31 AV618931 AV618931
25 53.6 5.2 1101 229 CNS001T2 AL078714 Drosophila
26 53.6 5.2 866 146 BF267363 BF267363 HV_CEA001
27 53.4 5.2 796 31 AV704385 AV704385 AV704385
28 53.4 5.2 797 106 AL534423 AL534423 AL534423
29 53.2 5.2 970 229 CNS0182A AL108796 Drosophila
30 53 5.2 839 229 CNS011KX AL100395 Drosophila
31 53 5.2 1092 230 CNS020K7 AL173696 Tetradon
32 53 5.2 1225 229 CNS0161D AL106171 Drosophila
33 52.8 5.1 633 105 AL513979 AL513979 AL513979
34 52.8 5.1 1029 230 CNS01ZGM AL174271 Tetradon
35 52.8 5.1 1101 229 CNS003B4 AL064082 Drosophila
36 52.6 5.1 767 229 CNS00AOX AL055924 Drosophila
37 52.6 5.1 860 170 BF867222 BF867222 963087E12
38 52.4 5.1 576 230 CNS035N7 AL028940 Tetradon
39 52.4 5.1 1101 229 CNS0064G AL062781 Drosophila
40 52.4 5.1 1101 229 CNS008X3 AL052344 Drosophila
41 52.2 5.1 545 115 AW447058 AW447058 88154 MAR
42 52 5.1 524 229 CNS01U90 AL167541 Tetradon
43 51.8 5.0 670 150 BF617942 BF617942 HVSMEC000
44 51.8 5.0 928 229 CNS00DKY AL071865 Drosophila
45 51.8 5.0 1101 229 CNS01219 AL101595 Drosophila
```

## ALIGNMENTS

```
RESULT 1
CNS012FZ 1094 bp DNA GSS 26-JUL-1999
LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC
DEFINITION BACN07B02 of DrosBAC library from Drosophila melanogaster (fruit
AL101513
AL101513 genomic survey sequence.
AL101513.1 GI:5613124
GSS.
fruit fly.
Plasmid Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1094)
REFERENCE
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.edi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBel0BAC11.
```

## FEATURES

## source

```
1.1094
/organism="Drosophila melanogaster"
/plasmid="pBel0BAC11"
/db_xref="taxon:7227"
/clone_lib="DrosBAC"
/clone="BACN07B02"
/note="end : T7"
BASE COUNT 466 a 151 c 25 g 256 t 196 others
ORIGIN
```

Query Match

6.8%; Score 70.4; DB 229; Length 1094.

Best Local Similarity 36.1%; Pred. No. 3.3e-06; Matches 208; Conservative 90; Mismatches 273; Indels 5; Gaps 2;

```
QY 450 TAAACAGCCGATATATATGAGCAAGCATGATGATACCAAGAGGTTAAACAGA 509
||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 524 TAAAAATTCWMTTAAACVAMCAACACCCCAACTTTCACAAAAATTTAAAAAW 583

QY 510 TGTCTATAGTTAAGACATGCACGACGAGGTATCTAGTAGAGCTACGTTGAT 569
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 584 YWCTTAAATTCACCAACCAAAATTTWOMATTTWYCCAAVCAABCTYCATTCAMW 643

QY 570 AGATGATCCACATATATCTACAGTATCAAGATGAGAAAAACCTTAAAGTAAT 629
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 644 TTTTAAAAAAAYTWCACAAAAATTTAATTTAATTAATTAATTAATTAATTAAT 703

QY 630 GTCATAGCTGGGATATATATATTTATTCAGCAATTTATCAAGATGATCAACCA 689
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 704 TTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 761

QY 690 TTGATATATATCTGAAATCCAAATTTTACAGCAATCAACATATGTAATTTAT 749
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 762 TTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 821

QY 750 CTGATAGCTGTTGATTTAACACATCTTCACATTTATTAAGAGAGTTGGAAC 809
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 822 TATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 881

QY 810 TTTGCTAGCTGAGGACGCTCCACACTACTCTGCAATTCCTCATTTATATAGTA 869
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 882 WTTTAAVCCYTTAAATAAAACACGAYCCTTAATMTAACTHTTTAATTAATTAAT 941

QY 870 ATTATCTCTATATATGCCCCGAAATTAATTTGGGATCTGTAATATATATAT 929
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 942 TTTTAAATTAATTAATTTTCYCTTCAMAAAAAATTAATTAATTAATTAATTAAT 1001

QY 930 AACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 989
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1002 AAAAAAAATWMAA---AAWMTAAWMAWCAAAAAAATTTCTTTTTCACAAW 1058

QY 990 ATTAAATCAAAATGTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1025
| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1059 WHAAAAAAHWAACMYTTTCMATCTTTTMAAA 1094
```

```
RESULT 2
AV426745 420 bp mRNA EST 23-MAY-2000
LOCUS AV426745 Lotus japonicus young plants (two-week old) Lotus
DEFINITION japonicus cDNA clone MM071b02_r 5', mRNA sequence.
ACCESSION AV426745
VERSION AV426745.1 GI:7785990
KEYWORDS EST.
SOURCE Lotus japonicus.
ORGANISM Lotus japonicus.
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
Fabales; Papilionaceae; Papilionoideae; Lotus.
1 (bases 1 to 420)
REFERENCE
AUTHORS Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
TITLE Generation of 7137 non-redundant expressed sequence tags from a
legume, Lotus japonicus
JOURNAL DNA Res. 7 (2), 127-130 (2000)
MEDLINE 20277479
COMMENT Contact: Yasukazu Nakamura
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: ynakam@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
```

```
FEATURES
source
1.420
/organism="Lotus japonicus"
/db_xref="taxon:34305"
/clone="MM071b02_r"
Location/Qualifiers
```

```

/c/clone.lib="rotus japonicus young plants (two-week old)"
/dev_stage="young plants (two-week old)"
/note="Vector: pbluescriptII SK-. Site_1: EcoRI; Site_2:
XhoI; isolate=Miyakojima M3-20"
BASE COUNT
ORIGIN
98 a 72 c 132 g 118 t

```

Query Match:	6.6%	Score 68.2:	DB 30	Length 420:
Best Local Similarity:	55.8%	Pred. No. 9.7e-06:		
Matches 130, Conservative	0;	Mismatches 103;	Indels 0;	Gaps 0

[illegible]

RESULT	3
CNS00396/c	
LOCUS	
DEFINITION	CNS00396 . 1101 bp DNA GSS 03-JUN-1999
	Drosophila melanogaster genome survey sequence TENG end of BAC #
	BACR08K10 of RPc1-98 library from Drosophila melanogaster (fruit
	fly), genomic survey sequence.
Accession	

KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL
fruit fly.						
Genoscope.						
Drosophila melanogaster						
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.						
1 (bases 1 to 1101)						
Genoscope.						
Direct Submission						
Submitted (02-JUN-1999)						
Genoscope - Centre National de Sequen						

**JOURNAL**  
Submitted (02-JUN-1999) Genoscope - Centre National de Séquençage  
BP 191 91006 EVRY cedex - FRANCE (E-mail : [segrete@genoscope.cns.fr](mailto:segrete@genoscope.cns.fr))  
- Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
**COMMENT**  
Determination of this BAC-end sequence was carried out as part of a

The BDBP is constructed using a physical map of the *Drosophila melanogaster* genome using these BACs. For further information please see <http://www.fruitfly.org> The BDBP *Drosophila melanogaster* BAC library was prepared by Kazuo Oseigawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCR-98 and was constructed by partial EcoRI digestion of *Drosophila* DNA provided by the BDBP from the isogenic strain y2; cn bw sp, the same strain used for the BDBP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

BASE COUNT	201 a	64 c	131 g	202 t	503 others
ORIGIN	/organism="Drosophila melanogaster" /db_xref="taxon:7227" /clone_lib="RPCT-98" /clone="BACROBK10" /note="end : TE93"				

Query Match	6.1%;	Score 63.2;	DB 229;	Length 1101;
Best Local Similarity	17.4%;	Pred. No. 0.00015;		
Matches 110;	Conservative 276;	Mismatches 244;	Indels 4;	Gaps 2;

OY	189	TATTACGTTAACAGCATTACTACTCTCTTTTGGTTCGAACACCAACCCGCTTTAAACAAATAT	24 8
Db	1100	KARRRGDDTWMRDTRKDDMDMTKMTTWMKDRADDRRNAGADRNAMDGAGTWTATW	104 0
OY	249	TGATTTAGATAGCAAAAGATAGTATACGCAACAAAGACAATTCGTTAGACACATGG	308 0
Db	1040	WWWATTTTWMDDKMMWWATAAKTDFTAWTWRTWRAADMGRDRGACKRRDRAATDADGA	98 13
OY	309	TGATGAGACGTTACGGGCTCACGGTATTGAGGTTGAATG--CGTTGATGATGAACGGCA	366 6
Db	980	GRDGRRRKRRKRRKDDDDKKGGKKKKAKAKAKMTATKMWDDMDKMKMGAKDKRA	92 11
OY	367	TCACATATATACCAAGACTTTTAAAGCAAAAGCAAGCAACGACCCACAATTTACAGTG	426 6
Db	920	DDDDGAGGKDDGKGAADDGTDGTRDDDKDDKMDMDKKGITMGDAITWMAATDMMW	86 11
OY	427	AAAGTATCTGCAAGTTTATGATGTTAAACAACCGCATATATGACACAAGTCAATGGATT	486 6
Db	860	WADDWMTWMDAAADDWMDAWMKWMDMAWMTGARTADRPWGRPAKRGARRRD	80 11
OY	487	ACTAACAAGAAGGTTTAAACAAGATGCTCATAGTATAGATACATGACACGACGAGTGTA	54 6
Db	800	KRRADKRDADDDDDAATWTWTTTTTTTRTDIDPMKKMTDTWTRAAADRTWDRDDDDDR	74 11
OY	547	ACT--GGAAGACGTACAGTTGATTTAGATGATGCACCAATATCTACACGATTTCAAGATG	604 6
Db	740	AGTGRKKRRRTTKRRMRROTRMDADADADTARDRRRRDDGDAGKRGKTRRRRD	68 11
OY	605	GAAAAAACCCATATAAGTATATGTCCTAAGTCGGGAATATTCATTTTAACTACAGCAA	664 6
Db	680	RATMDRTDAMWADAMWTTDTDDMDKRRRRKKRRRRRTTARAAAMWMTWKAMDWA	62 11
OY	665	TTTATCAAGATGATTCACACCAATTTGGATATATCTAGTAATAATCCAAATTTTAAACGA	724 6
Db	620	KWDKTRADRRDRRAADDTWIDARADADMAKARWRARRDRRAARADRRWTGKTTTA	50 11
OY	725	ATCAACAACATATTTGAATTAATTTACTTGAAGTCCTTGATTTTAAACAACATCTTCACA	784 6
Db	560	TWTTWAAAAAAMWMAWATATATTTWTTTWTWTTTTTTTTTTTAAWMAWMTATW	50 11
OY	785	ATTATATATAAAGAGGAGTTGGAAGCTTGTCTAGT	81 8
Db	500	AAWTTAAAAAAMAAATTTTTTTTTTTTTTWTWMT	46 7

RESULT	4
LOCUS	BE920828
DEFINITION	BE920828 469 bp mRNA EST
ACCESSION	EST442597 potato leaves and petioles Solanum tuberosum cDNA clone.
VERSION	CSTB1019 5' sequence, mRNA sequence.
KEYWORDS	BE920828 BE920828.1 GI:10446904
SOURCE	EST.
ORGANISM	potato. Solanum tuberosum
REFERENCE	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasteridae I; Solanales; Solanales; Solanales; Solanum.
AUTHORS	I (bases 1 to 469) van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J., Uiterback,T.T., Hansen,C.L., Doan,B., Bougri,O., Buell,C.R., Ronning, C.M., Fry,W.E., Tanksley,S.D. and Baker,B.
TITLE	Generation of ESTs from potato leaves and petioles
JOURNAL	Unpublished (2000)
COMMENT	Contact: Cathy Ronning The Institute for Genomic Research For clone request: please contact Research Genetics, Libraries



QY	911	GAAATTATCAATTTATCATCAACCAATGATGTGATT	945
		: :     :   :	
Db	1032	TAATATATWATAWWWTAWATATATANAWWAATTAWTT	1066



```

RESULT 8
CNS01807 1101 bp DNA GSS 26-JUL-1999
LOCUS CNS01807/c
DEFINITION Drosophila melanogaster genome survey sequence sp6 end of BAC
BACN37F07 of DrosBAC library from Drosophila melanogaster (fruit
fly); genomic survey sequence.
ACCESSION AL108721 GI:5629025
VERSION AL108721.1
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
AUTHORS BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
JOURNAL - Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk/. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.
FEATURES
Source location/Qualifiers
1. 1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="DrosBAC"
/clone="BACN37F07"
/notes="end : Spe"
BASE COUNT 179 a 70 c 85 g 448 t 319 others
ORIGIN
Query Match 5.9%; Score 61.2; DB 229; Length 1101;
Best Local Similarity 42.5%; Pred. No. 0.00044;
Matches 182; Conservative 21; Mismatches 221; Indels 4; Gaps 1;
QY 371 AATTATACCAAGACTTTTAAAGCAAAAGCAAGCACTGCCAATATTACAGTGAAG 430
| : | | | : | | | | | | | | | | | | | | | | | | | | | | | |
Db 562 AAAAAATAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 503
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 431 TATCTGCACTTATAGATGTAACAACAGCAATGATAATGACAAAGTCAATGATTACTA 490
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 502 TAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 443
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 491 ACAAGAGCTTAACAAGATGCTATAGTTAAGTACATCGACAGCGAGCTGTTACTG 550
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 442 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 383
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 551 GAAGACGTACAGTTGATTTGATGATCCACATATTACTACAGTATTCAGATGAAAA 610
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 382 AATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 323
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 611 ACCCTATTAAGTAATTTCTTAAGTCTGGAATATTTCATTTTAAATCAGCAATTTTATC 670
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 322 ATATATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 267
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 671 AAGATGAATCAACACCAATTTGATATATCTAGTCAAAATCCAAATTTTACAGCAATTA 730
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 266 AATTTTATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 207
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 731 CACATATTGAATTAATTTACTGAGCTGTGATTTTACACCAATTTCTTCAATTTAT 790
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 206 WAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 147
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

```

QY 791 AAAAAAGA 798
| : | : | : |
Db 146 AAAAAAAA 139
| : | : | : |
RESULT 9
CNS001J7 1101 bp DNA GSS 03-JUN-1999
LOCUS CNS001J7
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC:
BACR30P10 of RPCI-98 library from Drosophila melanogaster (fruit
fly); genomic survey sequence.
ACCESSION AL068307 GI:4958538
VERSION AL068307.1
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
AUTHORS BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
JOURNAL - Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazuhiro Osoegawa and
Aaron Mammoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
PI and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES
Source location/Qualifiers
1. 1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RPCI-98"
/clone="BACR30P10"
/notes="end : T7"
BASE COUNT 507 a 148 c 112 g 171 t 163 others
ORIGIN
Query Match 5.9%; Score 60.2; DB 229; Length 1101;
Best Local Similarity 37.0%; Pred. No. 0.00076;
Matches 160; Conservative 54; Mismatches 219; Indels 0; Gaps 0;
QY 369 ACAATTATACCAACACTTTTAAAGCAAAAGCAACCACTGCCAATAATTACAGTGA 428
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 121 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 180
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 429 AGTATCTGCAAGTTTATGATGTAACAACGCAATGATATGACAAAGCAATGATTTAC 488
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 181 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 240
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 489 TAACAAGAGGTTTAACAAGATGCTTAAGTTAAGCATCGACAGCAGCAGCTTTAAC 548
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 241 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 300
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 549 TCGAAGACGTACAGTTGATTTAGATGATCCACATATATCTACAGTTTCAAGTGA 608
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 301 WAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 360
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 609 AAACCTATTAAGTAATTTCTTAAGTCTGGAATATTTCATTTTAAATCAGCAATTTTA 668
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```









**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 7, 2001, 10:38:44 ; Search time 66.69 Seconds  
(without alignments)  
2694.050 Million cell updates/sec

Title: US-09-376-633-1  
Perfect score: 1029  
Sequence: 1 ATGAGTTATGCGATTCACT.....TAACTTTACGAAGAAGTGA 1029

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 302621 segs, 87301344 residues

Total number of hits satisfying chosen parameters: 605242

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/prodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/prodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/prodata/2/ina/5A\_COMB.seq:\*  
4: /cgn2\_6/prodata/2/ina/5B\_COMB.seq:\*  
5: /cgn2\_6/prodata/2/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/prodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1029	100.0	1029	US-08-978-456-1	Sequence 1, Appl
2	1029	100.0	3336	US-08-927-554-2	Sequence 7, Appl
3	1029	100.0	3336	US-08-978-456-7	Sequence 7, Appl
4	1029	100.0	3336	US-09-225-967-7	Sequence 7, Appl
5	352.2	34.2	1269	US-08-978-456-3	Sequence 3, Appl
6	112.6	10.9	3417	US-08-978-458-7	Sequence 7, Appl
7	112.6	10.9	3417	US-08-978-454-7	Sequence 7, Appl
8	112.6	10.9	3417	US-09-385-288-7	Sequence 7, Appl
9	76.2	7.4	1230	US-08-741-327E-14	Sequence 14, Appl
10	55.6	5.4	7218	US-08-232-463-14	Sequence 14, Appl
11	51.2	5.0	3095	5231168-1	Patent No. 5231168
12	47.8	4.6	9636	US-08-323-170B-1	Sequence 11, Appl
13	47.2	4.6	1826	US-09-286-691-11	Sequence 11, Appl
14	46.4	4.5	5589	US-08-465-795-1	Sequence 13, Appl
15	43.4	4.2	19124	US-08-487-826B-13	Sequence 13, Appl
16	41.6	4.0	4766	PCT-US93-07261-10	Sequence 10, Appl
17	40.6	3.9	6152	US-08-973-462-1	Sequence 1, Appl
18	40.4	3.9	665	US-08-883-795A-36	Sequence 36, Appl
19	39.2	3.8	4430	US-08-918-914-2	Sequence 2, Appl
20	38.8	3.8	1186	US-08-731-722-5	Sequence 5, Appl
21	38.8	3.8	3781	US-08-688-988-5	Sequence 5, Appl
22	38.8	3.8	19124	US-08-487-826B-13	Sequence 13, Appl
23	38.6	3.7	2817	PCT-US93-05944-1	Sequence 1, Appl
24	38.2	3.7	2520	US-08-450-351-1	Sequence 3, Appl
25	38.2	3.7	4248	US-08-678-614-1	Sequence 1, Appl
26	38.2	3.7	1939	US-07-715-731B-2	Sequence 2, Appl
27	38	3.7	1939	US-07-715-731B-2	Sequence 2, Appl

28	38	3.7	10968	2	US-08-680-327-2	Sequence 2, Appl
29	37.8	3.7	2058	2	US-08-749-391-1	Sequence 1, Appl
30	37.8	3.7	2058	4	US-09-390-200-1	Sequence 1, Appl
31	37.2	3.6	1258	1	US-08-335-518-1	Sequence 1, Appl
32	37.2	3.6	1258	4	US-08-988-054-1	Sequence 1, Appl
33	37	3.6	2237	1	US-08-487-1355-1	Sequence 1, Appl
34	37	3.6	2237	2	US-08-915-972A-1	Sequence 1, Appl
35	37	3.6	2237	2	US-09-177-909-1	Sequence 1, Appl
36	36.6	3.6	446	4	US-09-097-541-1	Sequence 1, Appl
37	36.6	3.6	3926	2	US-08-731-722-1	Sequence 1, Appl
38	36.6	3.6	3926	2	US-08-731-722-2	Sequence 2, Appl
39	36.6	3.6	6216	4	US-09-213-053-1	Sequence 1, Appl
40	36.6	3.6	8920	2	US-08-446-855A-1	Sequence 1, Appl
41	36.6	3.6	8920	4	US-09-150-741-1	Sequence 1, Appl
42	36.6	3.6	56516	2	US-08-996-306-1	Sequence 1, Appl
43	36.4	3.5	3850	3	US-08-329-799-34	Sequence 34, Appl
44	36	3.5	3659	1	US-08-320-151-72	Sequence 72, Appl
45	36	3.5	3659	1	US-08-303-124-12	Sequence 12, Appl

## ALIGNMENTS

```
RESULT 1
US-08-978-456-1
; Sequence 1, Application US/08978456
; Patent No. 6010881
; GENERAL INFORMATION:
; APPLICANT: Palmer, Leslie M.
; TITLE OF INVENTION: NO. 6010881el ribc
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Dechart Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: US
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/978,456
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, Todd O
; REGISTRATION NUMBER: 28,354
; REFERENCE/DOCKET NUMBER: P50444-9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2252
; TELEFAX: 215-994-2222
; TELEX:
; INFORMATION FOR SEQ. ID NO.: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1029 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: Linear
; US-08-978-456-1

Query Match 100.0%; Score 1029; DB 3; Length 1029;
Best Local Similarity 100.0%; Pred. No. 3.7e+246;
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 ATGAGTTATGCGATTCACTTCAAAATATGCTCAAGCTCAACAGGTGTTAATCCACC 60
|||||
```

Db	1	ATGATTATGCGATTCACAACTTGCAAATATGTCACAAAGTCACAAACAGGTATTAATCACACC	60
QY	61	GTTCGGCCCTGTGTAGTAAATGAAGTAGGATTTGTGTTATGTGTGCACACTTGGAGAAA	120
Db	61	GTTCGGCCCTGTGTAGTAAATGAAGTAGGATTTGTGTTATGTGTGCACACTTGGAGAAA	120
QY	121	GCTACCAAGATGGGGAGTTCCAAGCACTGATATGCGCACCAACAAATCTGAAGTGGC	180
Db	121	GCTACCAAGATGGGGAGTTCCAAGCACTGATATGCGCACCAACAAATCTGAAGTGGC	180
QY	181	ACGATTTATATATACGTTAGAGCCATGTAGTCATTTTGTTCACACACCCTGTGTTAAC	240
Db	181	ACGATTTATATATACGTTAGAGCCATGTAGTCATTTTGTTCACACACCCTGTGTTAAC	240
QY	241	AAATTTATGATTTGTAAAGATACAAAAGTATGTAACGCAACAAAGACATTCGTTAGAC	300
Db	241	AAATTTATGATTTGTAAAGATACAAAAGTATGTAACGCAACAAAGACATTCGTTAGAC	300
QY	301	ACACATGCTATGAGAGCTTACGGGCTCACGGTATGAGGTTGAATGCGTGTATATGTA	360
Db	301	ACACATGCTATGAGAGCTTACGGGCTCACGGTATGAGGTTGAATGCGTGTATATGTA	360
QY	361	CGGGCATCACAAATTTATCCAAAGACTTTTTTAAAGCAAAAGCAAGCACTGCGACAAAT	420
Db	361	CGGGCATCACAAATTTATCCAAAGACTTTTTTAAAGCAAAAGCAAGCACTGCGACAAAT	420
QY	421	ACAGTGAAGATATCTGCAGTTAGTATGTTAGTAAACAAGCGATGATATAGGACAAAGTCAA	480
Db	421	ACAGTGAAGATATCTGCAGTTAGTATGTTAGTAAACAAGCGATGATATAGGACAAAGTCAA	480
QY	481	TGGATTTCTAAACAAAGGTTAAACAAGATGCTATATACGTTAAACATGACACAGACGA	540
Db	481	TGGATTTCTAAACAAAGGTTAAACAAGATGCTATATACGTTAAACATGACACAGACGA	540
QY	541	GTGTTAACTGGAAGACCTACAGTTGTAATGATGATCCAAATATCTACAGTATTCAA	600
Db	541	GTGTTAACTGGAAGACCTACAGTTGTAATGATGATCCAAATATCTACAGTATTCAA	600
QY	601	GATGGAAGAAAAACCTATTAAGATATATGTTCTAGTCTGGGAATTTGATTTAATCAG	660
Db	601	GATGGAAGAAAAACCTATTAAGATATATGTTCTAGTCTGGGAATTTGATTTAATCAG	660
QY	661	CAATTTATCAAGTGAATGAAACCAATTTGGATATATCTGAAATTCCAATTTTACA	720
Db	661	CAATTTATCAAGTGAATGAAACCAATTTGGATATATCTGAAATTCCAATTTTACA	720
QY	721	AGCATCAACACATATTTGAATTTATTTACTTGAAGCTTGGATTTTAAACAACATTTCTT	780
Db	721	AGCATCAACACATATTTGAATTTATTTACTTGAAGCTTGGATTTTAAACAACATTTCTT	780
QY	781	CACAAATTTATATTAAGAGAGGATTTGAACTTTGCTAGTGCAGGACAGTCCAACTACT	840
Db	781	CACAAATTTATATTAAGAGAGGATTTGAACTTTGCTAGTGCAGGACAGTCCAACTACT	840
QY	841	TCAAAATTCCTCCATATTTATAGATGAATTTATCTCTATTTATGCCCGCAAAATTAAT	900
Db	841	TCAAAATTCCTCCATATTTATAGATGAATTTATCTCTATTTATGCCCGCAAAATTAAT	900
QY	901	GGCGGATCTGGAATTTATCAATTTTATCAAAACAATGATGTGATGATACAGATGGC	960
Db	901	GGCGGATCTGGAATTTATCAATTTTATCAAAACAATGATGTGATGATACAGATGGC	960
QY	961	AACCAATTTGAATTTGTTCTATTCGAGTTATTAATCAAAATGTTAAATTAACCTTTACA	1020
Db	961	AACCAATTTGAATTTGTTCTATTCGAGTTATTAATCAAAATGTTAAATTAACCTTTACA	1020
QY	1021	AAGAAGTGA 1029	
Db	1021	AAGAAGTGA 1029	



```

Db 241 AAAATTATGATGTGAAGATAGCAAAAAGTGTATACGCAACAAAAGCAATTCGTTAGC 300
Oy 301 ACACATGCTGATGAGACGTTACGGGCTCAGCGTATGAGGTGAGTTCGTTGATGAA 360
Db 301 ACACATGCTGATGAGACGTTACGGGCTCAGCGTATGAGGTGAGTTCGTTGATGAA 360
Oy 361 CGGCGATCACAATTTATACCAAGACCTTTTAAACCAAAAGCAAGCACTGCCAATTT 420
Db 361 CGGCGATCACAATTTATACCAAGACCTTTTAAACCAAAAGCAAGCACTGCCAATTT 420
Oy 421 ACAGTGAAGTATCTGCAATTTAGATAGTAACCAAGCAAGTATGATGCAAGCAAGTCA 480
Db 421 ACAGTGAAGTATCTGCAATTTAGATAGTAACCAAGCAAGTATGATGCAAGCAAGTCA 480
Oy 481 TGGATTACTAACCAAGAGTTAAACAAGATGCTATAGTTAAGACATCGACAGCA 540
Db 481 TGGATTACTAACCAAGAGTTAAACAAGATGCTATAGTTAAGACATCGACAGCA 540
Oy 541 GTGTTAACTGGAAGACGTACAGTTGAATGATGATCCACAATATCTACACGTTTCAA 600
Db 541 GTGTTAACTGGAAGACGTACAGTTGAATGATGATCCACAATATCTACACGTTTCAA 600
Oy 601 GATGGAAGAAACCCCTATAAAGTAATTTGCTAAGCTGGGAATATCTATTTAATCAG 660
Db 601 GATGGAAGAAACCCCTATAAAGTAATTTGCTAAGCTGGGAATATCTATTTAATCAG 660
Oy 661 CAATTTATCAAGATGATCAACACCAATTTGATATATCTGTAATTCGAATTTTACA 720
Db 661 CAATTTATCAAGATGATCAACACCAATTTGATATATCTGTAATTTTACA 720
Oy 721 AGCAATCAACACATATTTGAAATTTATTTACTGTAAGCTGTGTGATTTAACAACTTCT 780
Db 721 AGCAATCAACACATATTTGAAATTTATTTACTGTAAGCTGTGTGATTTAACAACTTCT 780
Oy 781 CACAATTTATTAAGAGAGAGTGGAACTTGTGCTGAGGAGGCTCCAACTACT 840
Db 781 CACAATTTATTAAGAGAGAGTGGAACTTGTGCTGAGGAGGCTCCAACTACT 840
Oy 841 TCAGAAATTCATCTATATATATAGATTAATTTCTCTATTTATGAGCCGCAATTAAT 900
Db 841 TCAGAAATTCATCTATATATATAGATTAATTTCTCTATTTATGAGCCGCAATTAAT 900
Oy 901 GCGGATCTGGAATTTATCAATTTTATCAAAACAATGATGATGATGATGATGATGATG 960
Db 901 GCGGATCTGGAATTTATCAATTTTATCAAAACAATGATGATGATGATGATGATGATG 960
Oy 961 AACCAATTTGAATTTGCTCCTCCAGGTTTAAATCAAAANTGTAATTAATTAATTCGA 1020
Db 961 AACCAATTTGAATTTGCTCCTCCAGGTTTAAATCAAAANTGTAATTAATTAATTCGA 1020
Oy 1021 AAGAAAGTGA 1029
Db 1021 AAGAAAGTGA 1029

```

RESULT 3  
US-08-978-456-7  
Sequence 7, Application US/08978456  
Patent No. 6010861

GENERAL INFORMATION:  
APPLICANT: Palmer, Leslie M.  
TITLE OF INVENTION: NO. 6010861el r1bG  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dechert Price & Rhoads  
STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre  
CITY: Philadelphia  
STATE: PA  
COUNTRY: US  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible

```

; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/978,456
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, Todd O
; REGISTRATION NUMBER: 28,354
; REFERENCE/DOCKET NUMBER: P50444-9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2252
; TELEFAX: 215-994-2222
;
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3336 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
US-08-978-456-7

Query Match 100.0%; Score 1029; DB 3; Length 3336;
Best Local Similarity 100.0%; Pred. No. 5.3e-246;
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ATGGATTATGCGATTCACTTGCACAAATATGATCAAGTCAACAGGTGTTAATCCACC 60
Db 1 ATGGATTATGCGATTCACTTGCACAAATATGATCAAGTCAACAGGTGTTAATCCACC 60
Oy 61 GTTGCCCTCTGTGACTATATGAAGTATGATTTGTTGTTGTTGTTGTTGTTGTTGTTG 120
Db 61 GTTGCCCTCTGTGACTATATGAAGTATGATTTGTTGTTGTTGTTGTTGTTGTTGTTG 120
Oy 121 GTGACAGAGATCGGAGGTTCAAGCACTTATATGATGATGATGATGATGATGATGATG 180
Db 121 GTGACAGAGATCGGAGGTTCAAGCACTTATATGATGATGATGATGATGATGATGATG 180
Oy 181 ACGATTTATATACGTTAGAGCCATGTATGATTTTGTTCACACACCCTGTGTTAC 240
Db 181 ACGATTTATATACGTTAGAGCCATGTATGATTTTGTTCACACACCCTGTGTTAC 240
Oy 241 AAAATTATGATGTAAGATAGCAAAAGTATATAGCAACAAAGCAATTTGTTAGAC 300
Db 241 AAAATTATGATGTAAGATAGCAAAAGTATATAGCAACAAAGCAATTTGTTAGAC 300
Oy 301 ACACATGCTGATGAGACGTTACGGGCTCAGCGTATGAGGTGAGTTCGTTGATGAA 360
Db 301 ACACATGCTGATGAGACGTTACGGGCTCAGCGTATGAGGTGAGTTCGTTGATGAA 360
Oy 361 CGGCGATCACAATTTATACCAAGACCTTTTAAACCAAAAGCAAGCACTGCCAATTT 420
Db 361 CGGCGATCACAATTTATACCAAGACCTTTTAAACCAAAAGCAAGCACTGCCAATTT 420
Oy 421 ACAGTGAAGTATCTGCAATTTAGATAGTAACCAAGCAAGTATGATGCAAGCAAGTCA 480
Db 421 ACAGTGAAGTATCTGCAATTTAGATAGTAACCAAGCAAGTATGATGCAAGCAAGTCA 480
Oy 481 TGGATTACTAACCAAGAGTTAAACAAGATGCTATAGTTAAGACATCGACAGCA 540
Db 481 TGGATTACTAACCAAGAGTTAAACAAGATGCTATAGTTAAGACATCGACAGCA 540
Oy 541 GTGTTAACTGGAAGACGTACAGTTGAATGATGATCCACAATATCTACACGTTTCAA 600
Db 541 GTGTTAACTGGAAGACGTACAGTTGAATGATGATCCACAATATCTACACGTTTCAA 600
Oy 601 GATGGAAGAAACCCCTATAAAGTAATTTGCTAAGCTGGGAATATCTATTTAATCAG 660
Db 601 GATGGAAGAAACCCCTATAAAGTAATTTGCTAAGCTGGGAATATCTATTTAATCAG 660

```





TELEX:  
: INFORMATION FOR SEQ ID NO: 7:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 3417 base pairs  
: TYPE: nucleic acid  
: STRANDEDNESS: double  
: TOPOLOGY: linear  
US-08-978-458-7

Query Match 10.9%; Score 112.6; DB 2; Length 3417;  
Best Local Similarity 49.2%; Pred. No. 4.1e-19;  
Matches 386; Conservative 0; Mismatches 389; Indels 10; Gaps 3;

```
OY 1 ATGATTTATGCGATTCACACTTGCACAAATATGTACAAAGTCAAGCTTTATCCACCC 60
DB 19 ATGAATTTAGCAATAAACTGGACAAAAAGGGCGTGTAGCTCAATCCCAATCCTTANG 78
OY 61 GTTGGCGCTGTTGTAGTTAATGAAGTATGTTGTTGTTGTTGTTGTTGTTGTTGTT 120
DB 79 GTTGGCGCAATTTATGTAAGATATATCAATTCGCAAGGTTATCATGATTTT 138
OY 121 GGTGACAAAGCATCGGAGGTTCAAGCACTTGATTCGACAAACAAATGCTGAAGTGGC 180
DB 139 GGTGGCCCAATGCTGAGAGAAATGCTTTAAACTGTGAAATCCCTGTGAGGCG 198
OY 181 ACGATTTATTTACGTTAGACCATGTACTCATTTTGTTCACACACCCCTGTGTTAAC 240
DB 199 ACGCTTTATGTAACACTTGACCTGTTGCTACTTCGGGAAAAACACCTCCTGTATAGAT 258
OY 241 AAAATTTATGATTTGTAAGTACCAAAAGTATACGCAACAAAGA-----CAAT 292
DB 259 GCTATTAATGATAGTGTATTTACAAAGATGATTCATTTGGAAGCCTGACGTATTCCTATT 318
OY 293 CCTTAGACACACATGTCATGAGAGCTTACGGGCTCACCGTATTTGAGTTGAGCTTG 352
DB 319 GTATCTGGAAGAAAGAGTAAGTACTGTGAGAAATTAATCTTCAAGTTACTGTGGAAT 378
OY 353 ATGATGAACGGGCATCACAATTAATACCAAGACTTTTAAAGCA-AAAGCAAGCAACTG 411
DB 379 TTAGAAATGATGTCCTTAATTAATTAAGTTTAAAGTATTTACCAAGCATGTA 438
OY 412 CCACAATTAACGTGAAGTATCTGCAATTTAGATGTAACCAAGGCAATATATGA 471
DB 439 CCTATGTTTATGAAATATGCAATGTCATATGAGATGAAATATGCCCTAAACCAAT 498
OY 472 CAAAGTCATGATTTACTTAACAAGAGTTTAAACAAGATGCTATTAAGTAAACATGCA 531
DB 499 CAATGCAATGATTTACTTAAGAGAGCAAGAAAGCATGTCATCAGTTAGACACTAT 558
OY 532 CACGACGCACTTTAACTGGAAGAGTACAGTTGAATTAAGATTCACCAATATACTACA 591
DB 559 GTTATGCAATTTATGTCGAGTCAATCTGTTATTCAGAGCATCTTGGTGAACATGT 618
OY 592 CATTATTCAGATGAAACCACTTATTAAGTAATATTTGTAAGTGGGAATATCAT 651
DB 619 ACATTGAGAGGAGGAAATATCCTATCCGTATCATATGCGATACACATTTACGAATCCT 678
OY 652 TTTTA-ATCAGCAAAATTTATCAAGATGAATCAACACCAATTTGATATATCTGAAATCC 710
DB 679 CTACTCTTAATATGTAAGAAACAGCAATATTAATTAACCTTACATTCCTCCACTCTCT 738
OY 711 AAATTTAACACCAATCAACACATATTTGAAATTTATTTACTGGAAGCTGTGATTTAAC 770
DB 739 GAAGCAAAATTAATGAAGTATATCAAAATCATGCTGTGAATATCTTCCATTAAG 798
OY 771 AACAA 775
DB 799 AAAAA 803
```

RESULT 7  
US-08-978-454-7

```
: Sequence 7, Application US/08978454
: Patent No. 6017728
: GENERAL INFORMATION:
: APPLICANT: Palmer, Leslie M.
: APPLICANT: Redon, Jason C.
: APPLICANT: Warren, Richard L.
: APPLICANT: Kosmatka, Anna L.
: APPLICANT: Shilling, Lisa K.
: APPLICANT: Stodola, Robert K.
: APPLICANT: Knowles, David J. C.
: APPLICANT: Black, Michael T.
: APPLICANT: Hodgson, John E.
: APPLICANT: Nicholas, Richard O.
: TITLE OF INVENTION: ribh
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dechert, Price & Rhoads
: STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
: CITY: Philadelphia
: STATE: PA
: COUNTRY: USA
: ZIP: 19103-2793
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/978,454
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/911,503
: FILING DATE: 15-AUG-1997
: APPLICATION NUMBER: PCT/US97/14436
: FILING DATE: 15-AUG-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Dickinson, O. Todd
: REGISTRATION NUMBER: 28,354
: REFERENCE/DOCKET NUMBER: P50533-11
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 215/994-2252
: TELEFAX: 215/994-2222
: TELEX:
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3417 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
US-08-978-454-7
```

Query Match 10.9%; Score 112.6; DB 3; Length 3417;  
Best Local Similarity 49.2%; Pred. No. 4.1e-19;  
Matches 386; Conservative 0; Mismatches 389; Indels 10; Gaps 3;

```
OY 1 ATGATTTATGCGATTCACACTTGCACAAATATGTACAAAGTCAACAGCTTTATCCACCC 60
DB 19 ATGAATTTAGCAATAAACTGGACAAAAAGGGCGTGTAGCTCAATCCCAATCCTTANG 78
OY 61 GTTGGCGCTGTTGTAGTTAATGAAGTATGTTGTTGTTGTTGTTGTTGTTGTTGTT 120
DB 79 GTTGGCGCAATTTATGTAAGATATATCAATTCGCAAGGTTATCATGATTTT 138
OY 121 GGTGACAAAGCATCGGAGGTTCAAGCACTTGATTCGACAAACAAATGCTGAAGTGGC 180
DB 139 GGTGGCCCAATGCTGAGAGAAATGCTTTAAACTGTGAAATCCCTGTGAGGCG 198
OY 181 ACGATTTATTTACGTTAGACCATGTACTCATTTTGTTCACACACCCCTGTGTTAAC 240
DB 199 ACGCTTTATGTAACACTTGACCTGTTGCTACTTCGGGAAAAACACCTCCTGTATAGAT 258
```



OY	652	TTTA-ATCAGCAAAATTATTCAGATGATGAACACACCATTGGATTAATCTAACAAATCC	710
Db	679	CTTACTCTAAAATGTCTAAAAACAGCAAATGATATTAAAACTTACTTGCCACTTCTCT	738
OY	711	AAATTTAACCAACAATCAAAACACATAATTGAAATTATTTACTGAACTCTTGATTTAAC	770
Db	739	GAAAGCAAAAAATTAATGAAGCTATATCAAAATCATGCGTGTGAATVACTTTCATTAAG	798
OY	771	AACAA	775
Db	799	AAAA	803
 RESULT 9 US-08-741-327E-14 ; Sequence 14, Application US/08741327E ; Patent No. 5925354 ; GENERAL INFORMATION: APPLICANT: Fuller, et al. TITLE OF INVENTION: Riboflavin mutants as vaccines TITLE OF INVENTION: against Actinobacillus pleuropneumoniae NUMBER OF SEQUENCES: 17 CORRESPONDENCE ADDRESS: ADDRESSEE: G. Kenneth Smith STREET: 300 S. Wacker Drive CITY: Chicago STATE: IL COUNTRY: USA ZIP: 60606 COMPUTER READABLE FORM: MEDIUM TYPE: Floppy disk COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS SOFTWARE: ASCII CURRENT APPLICATION DATA: APPLICATION NUMBER: US/08/741,327E FILING DATE: October 28, 1996 CLASSIFICATION: 424 ATTORNEY/AGENT INFORMATION: NAME: G. Kenneth Smith REGISTRATION NUMBER: 43,135 REFERENCE/DOCKET NUMBER: 97704-E TELECOMMUNICATION INFORMATION: TELEPHONE: 312-913-0001 TELEFAX: 312-913-9808 INFORMATION FOR SEQ ID NO: 14: SEQUENCE CHARACTERISTICS: LENGTH: 1230 bases TYPE: nucleic acid STRANDEDNESS: single TOPOLOGY: linear US-08-741-327E-14			
 Query Match 7.4%; Score 76.2; DB 2; Length 1230; Best Local Similarity 47.6%; Pred. No. 3,1e-10; Matches 293; Conservative 0; Mismatches 313; Indels 9; Gaps 2			
OY	9	TGCAGATTCACCTTGGAAATATGTCAGCAAGCTCAAAACAGGTGTTAATCCCACCCGTGGCGC	68
Db	165	TGCCATTGCACCTGGCAAAAACAGGTTTAGGCTGGACGAATCCCAATCCCGCTGTGCGGTTG	224
OY	69	TGTTGTAGTTAATGAAGGTAGGATTTGGTATTTGGTGACACTTGAGAAAAGGTGACAA	128
Db	225	TGTAATTTGCAAAAACGGTGAATGCTTGCCGAAGTTTACCATGAAAGAAATTTGGTGANG	284
OY	129	GCATCGGAGGTTCACGACCTTGATATGGCACACAAAATGCTGAAGGTGCGACGATTTA	188
Db	285	GCATCGGAGACCTAATGCCGTTTTACATTTGTAAGGAAGATCTTCCGGGGCGACTGCTTA	344
OY	189	TATTAAGTTAAGGCATAGTACATTTTGGTTCACACACACACCCCTGTTTAAACAAAATTTAT	248
Db	345	TGTAAGGCTTGAGCCTTGTGTATCATACAGCGCGACGCGCCTTGTGGATTTATTAAT	404

```

QY 249 TGAATGTTAAGTATACCAAAAGTAGTATACGCAACAAAAGA--CAATTGCGTTAGACACACA 305
Db 445 TGAAGAGGACGATTAAAAAGATTATTATTCGCTTCACAGCATCCGAATCCTTTAGTACGACG 464
QY 306 TGGTG-----ATGAGACGTTACGGGCTACCGTGATTGAGGTGGAATGCGTTGATGTGA 359
Db 465 GCGGGGAGCAAAATCAGCTACGCGCCAGCGCGGTGGAAGTGTGGAGGTTACTCAAAAGA 524
QY 360 ACGGGCATACCAATTATATACCAAGACTTTTAAAGCAAAAGCAAAAGCAAGCACTGCCACAAT 419
Db 535 AGAATGTCATCGCTTAAACCCGATTTTTTTTCCACTATATTCATAAAGCTATCCGTAATGT 584
QY 420 TACAGTGAAGATATCTCGCAAGTTAGATGTTAAAGCAAGCAATGATATGACAAAATGCA 479
Db 585 GCTATATGAATAATTCGCATGACGCGAGACGAGCAAAATTTGCCAACCGGTAGCGCGCAATCCAA 644
QY 480 ATGATTTACTTACCAAAAGAGGTTAAACAAGATGTTCTATATAGTTAAGACATGCACGACGC 539
Db 645 ATGATTTACCGGTGAATTCGCAAGAGCAAGAGTGTCAGCAAAACAGTCATCAATATATGTC 704
QY 540 AGCTGTACTCGAAGACGCTACAGTTGATTTGATTTGATTCACCAATATATCTACAGCTATTCA 599
Db 705 GATTATGCTCGGTGTAGATTACGCTACTTGGCCGATTAACCGATGTTAATATGCGGAATGCC 764
QY 600 AGATGCAAAAAAACCC 614
Db 765 GAATCGCAAAACACCC 779

RESULT 10
US-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ. ID NO.: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
type: nucleic acid

```



NAME: Quine, Jonathan A.  
REGISTRATION NUMBER: P-41,261  
REFERENCE/DOCKET NUMBER: 015280-113100US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9636 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 149..9556  
US-08-323-1708-1

Query Match 4.6%; Score 47.8; DB 1; Length 9636;  
Best Local Similarity 49.0%; Pred. No. 0.0063;  
Matches 127; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

QY 460 ATGATATGCAAGCAAGTCAATGATTTACTACAAGAGCTTAACCAAGATGCTCTAAG 519  
DB 6941 AGTGATTTGAAGATGATCAATATTTATTTATTTCAATTAAGATTAATAATTAAT 7000  
QY 520 TTAAGACATGACAGACAGCAGCTGTTAACTGGAACGATGATTAATGATGATCA 579  
DB 7001 TTAAGACCAAGATGATTTATTAATTAATGATGATTAATTAAGATTAATAATTA 7060  
QY 580 CAAATCTACAGTATTCAGATGAAAAAACCCTATAAAGTATTTGCTAAGTCT 639  
DB 7061 CAACATATATACAGATATCAATATTTACAGATGCTTACATTTCAAAAATTTAATTA 7120  
QY 640 GGGATATCTTAATTAAGCAATTTTCAAGATGATTAACCAATTTGGATATAT 699  
DB 7121 GGTATCTACCAATTAATTTTAAATATCTTATTTACAGCATATGCTAAGATGAT 7180  
QY 700 ACTGAATCCCAATTTA 718  
DB 7181 ACCTTAATCTATATTTA 7199

RESULT 13  
US-09-286-691-11  
Sequence 11, Application US/09286691  
Patent No. 6190189  
GENERAL INFORMATION:  
APPLICANT: Li, Xin-Liang  
APPLICANT: Ljungdahl, Lars G.  
APPLICANT: Chen, Huizhong  
TITLE OF INVENTION: Cellulases and Coding Sequences  
FILE REFERENCE: 42-96  
CURRENT APPLICATION NUMBER: US/09/286,691  
CURRENT FILING DATE: 1999-04-05  
EARLIER APPLICATION NUMBER: US 60/037,883  
EARLIER FILING DATE: 1996-10-04  
EARLIER APPLICATION NUMBER: PCT US97/18008  
EARLIER FILING DATE: 1997-10-03  
NUMBER OF SEQ ID NOS: 29  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 11  
LENGTH: 1826  
TYPE: DNA  
ORGANISM: Orpinomyces sp. PC-2  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (69)..(1481)  
US-09-286-691-11

Query Match 4.6%; Score 47.2; DB 4; Length 1826;

Best Local Similarity 46.8%; Pred. No. 0.0053;  
Matches 148; Conservative 0; Mismatches 168; Indels 0; Gaps 0;

QY 535 GACGAGTGTATACAGGAGAGCTGACGTTGAATTAAGATGATCCCAATATCTACAGCT 594  
DB 1444 gatgggggtgtaaaatgtaattggtggtatatttaacttaacttaacttaactttaa 1503  
QY 595 ATTCAAGATGAAAAAACCCTATTAAGTATTTGCTTAAGCTGGGAATTTGATTTT 654  
DB 1504 caaacataataatattatttagtaataataaagaataaatttttaaaaaataatatt 1563  
QY 655 AATCAGCAATTTATCAAGATGATCAACACCAATTTGGATTTATACGAAATCCAAAT 714  
DB 1564 tatataatagttataataataataataataatagaattactatagttatataaga 1623  
QY 715 TTAACAGCAATCAACACATATTTGAATTTATTTACTTGAAGCTGTTGATTTAACAACA 774  
DB 1624 tataatacaatacaaaagtaaaaaaatttaaaattttagtattgataaaatttataa 1683  
QY 775 ATCTTCAATTTATTAATAAAGAGAGATTGGAAGCTTGCTAGTGAAGCAGTCCACC 834  
DB 1684 aagtttaataaagttaaaaaaataataataacatttggatgatttgcatacaaga 1743  
QY 835 ACTACTTCAGATTTCT 850  
DB 1744 aataataataaact 1759

RESULT 14  
US-08-465-795-1  
Sequence 1, Application US/08465795  
Patent No. 5589355  
GENERAL INFORMATION:  
APPLICANT: Koizumi, Satoshi  
APPLICANT: Yonetani, Yoshiyuki  
APPLICANT: Teshiba, Sadao  
TITLE OF INVENTION: A Process for Producing Riboflavin  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farbow, Garrett &  
SRETT: 1300 I Street, N.W., Suite 700  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/465,795  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/161,394  
FILING DATE: 06-DEC-1993  
APPLICATION NUMBER: JP 326578/1992  
FILING DATE: 07-DEC-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Gaybrick, Robert J.  
REGISTRATION NUMBER: 27,890  
REFERENCE/DOCKET NUMBER: 04853.0008-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5589 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear



MOLECULE TYPE: DNA (genomic)  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Corynebacterium ammoniagenes  
 US-08-465-795-1

Query Match 4.5%; Score 46.4; DB 1; Length 5589;  
 Best Local Similarity 47.1%; Pred. No. 0.012;  
 Matches 276; Conservative 0; Mismatches 301; Indels 9; Gaps 4;

QY 31 GTCACAGCTAAACAGCTGTTATTCACCCGCTGGCTGTTAGT---AATGAAGT 87  
 DB 672 GTGGCGGGAGCTACACCCGAGTCCACCGGTGGTGGTAATTATTCGACTCCGGT 731  
 QY 88 AGGATTTGGTATTTGGTGGCACTTGAAGAAAGGTGACAAGCATGGCGAGTCAAGCA 147  
 DB 732 GAGATTTGGGACCGGCTGACTCAGCCGGTGGCGGCTGCACGACAGAGTCCAACT 791  
 QY 148 CTTGATATGGACACAAATAATGCTGAAGTGCACGATTTATTTATTTAGAGCATGT 207  
 DB 792 CTAGCCGATCCCGGGGCAAGACCGAGGCGTACCCCGGTGGAGCTGGAGCCGTGC 851  
 QY 208 AGTCATTTGGTTCACACCACTCTGTAAACAAATTTATGATTTGAATAGCAAAA 267  
 DB 852 CGGCAATCCGGCGGACGGGACGCTGACGACGCTTTAATTTGAAGCCGGCATCAAGAT 911  
 QY 268 GTAGTATCCGACAAAGAC---AATCGTTAGACACATGTTGATGAGCTTACGG 324  
 DB 912 GTCTTTTCTTACACTCCGACCGAATCCAGTCGGCGGCGGGAACAGGTCCTTGTG 971  
 QY 325 GCTCAGCGTATTAGGT--TGAATGCTGTATGATGAAGCGGCAATTTATACCAAG 382  
 DB 972 GATGCGGATCAACGTCGTCGACGCTCCAGCCGGAGGGGATGCGGATGCCCTCAT 1031  
 QY 383 ACTTTTAAAGCAAAAGCA--ACCACTGCCCAATTTACAGTGAAGTATTCGCAAGT 441  
 DB 1032 CCGTGGTGAAGTCCGCTGCTGACGCTCCGACGCTGCTGTAATTTGCGCAACT 1091  
 QY 442 TTGATGTTAAACAGGATGATTAATGCAAAAGTCAATGATTTACTTACAAAGAGTT 501  
 DB 1092 ATGATGAGCTTTACTGAGCTGCGGATGACACGACGAGTGATCCTGGGAATGCG 1151  
 QY 502 AAACAGATGCTCTATAAGTAAAGATCGACAGCAGCAGTGTAACTGAGAGCACTACA 561  
 DB 1152 CGGACATGTCACAGCTGACCGCTGACGACGCGGATGCCATTTATCATCGGACCTGCACG 1211  
 QY 562 GTTGAATTAGATGATCCACAATATATCTACAGCTATTCAGATGGA 607  
 DB 1212 GCGTTGATGATATCATGCTACTGCGCCGCTATTCACAGACGCA 1257

RESULT 15  
 US-08-487-8268-13  
 Sequence 13, Application US/084878268  
 Patent No. 5993827

GENERAL INFORMATION:  
 APPLICANT: Sim, Kim L.  
 APPLICANT: Chitnis, Chetan  
 APPLICANT: Miller, Louis H.  
 APPLICANT: Peterson, David S.  
 APPLICANT: Su, Xin-zhaun  
 APPLICANT: Wellens, Thomas E.  
 TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX  
 AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS  
 NUMBER OF SEQUENCES: 45  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Knobbe Martens Olson & Bear  
 STREET: 620 Newport Center Drive 16th Floor  
 CITY: Newport Beach  
 STATE: California  
 COUNTRY: US

ZIP: 92660  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/487,826B  
 FILING DATE: 10-SEP-1993  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Israelson, Ned  
 REGISTRATION NUMBER: 29,655  
 REFERENCE/DOCKET NUMBER: NIH121.001CPI  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (619) 235-8550  
 TELEFAX: (619) 235-0176  
 INFORMATION FOR SEQ ID NO: 13:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 19124 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 US-08-487-8268-13

Query Match 4.2%; Score 43.4; DB 2; Length 19124;  
 Best Local Similarity 46.0%; Pred. No. 0.096;  
 Matches 188; Conservative 0; Mismatches 216; Indels 5; Gaps 1;

QY 390 TAAAGCAAAAGCAAGACCTGCCAAATTTACGTAAGTATCTGCAAGTTAGATGG 449  
 DB 15493 TAAAGCAAAAGCAAAAGCAAAAGCAAAATTTTAAAGAAATTAATTAATATATA 15552  
 QY 450 TAAACAGCGAATGATGACAGCAAGTCAATGATTTACTAACAAAGAGTTAAACAGA 509  
 DB 15553 AAATATTAATTTGATGAAATTAAGAAATGAAAGATTTACAAAAGAAATTTAAAA 15612  
 QY 510 TGCTATTAAGTTAAGACATGACACGACGCTGTTAACTGGAAGACGTCAGTTGAT 569  
 DB 15613 ATTTTATATAAAAAAAATGATTTATATAA---AAATAAAAACAAAGAGAAAA 15667  
 QY 570 AGATGATCCCAATATATCTACGTTATCAAGATGAGAAAAACCTTAAAGTAATATT 629  
 DB 15668 AAAAACAATTAAGCAAAAGCAAAATATATATATCAAAACAAAGAAAAAATATA 15727  
 QY 630 GTCCTAGCTGCGAATATTCATTTTAAATCAGCAATTTATCAAGATGAATCAACCAAT 689  
 DB 15728 TTAATAATTAATATATATCTATTAATAAAGAAATTTAAAGAAATTTAAAGAAAT 15787  
 QY 690 TTGATATATATAGTGAAGAAATCCAAATTTAAACAGCAATCAACACATATTGAATTTA 749  
 DB 15788 ATATACATTAATAAATAAATAATTTATTAATAAAGAAATTAATAAATAAATTTA 15847  
 QY 750 CTGAAGCTTTGATTTAACAACAATTTCTTCAAAATTTATATAAGA 798  
 DB 15848 ATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 15896

Search completed: May 7, 2001, 12:12:52  
 Job Time: 5648 sec

**THIS PAGE BLANK (USPTO)**